



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 123552

TO: Celine Qian
Location: REM-2C70/2A89
Art Unit: 1636
Friday, June 04, 2004

Case Serial Number: 09/515363

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Qian,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor
571-272-2507 Remsen E01 D86

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library Remsen Bldg.



Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches. Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.rnpm** and **.rnpn**. Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**.

Because they contain data that is confidential, the results of Pending database searches should not be left in the case .

077N: Ed Hart.

123552

Access DB#

SEARCH REQUEST FORM

Scientific and Technical Information Center

CRFE

Requester's Full Name Celine Qian Examiner # 78710 Date: 6/2/04
An Unit 1636 Phone Number 303-272-0777 Serial Number 09/515363
Mail Box and Bldg Room Location 2A89 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention Melanoma differentiation associated gene-5
Inventors (please provide full names): Paul Fisher et al.

Earliest Priority Filing Date: 2/29/2000

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search for SEQ ID NO: 1 & 2.

STAFF USE ONLY

Searcher _____

Searcher Phone # _____

Searcher Location _____

Date Searcher Requested 6/2/04

Date Submitted 6/4/04

Searcher Prep. Review Time _____

Final Prep Time _____

Final Time _____

Type of Search

NA Sequence (#) 1

AA Sequence (#) 1

Structure (#) _____

Bibliographic _____

Litigation _____

Full text _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN _____

Dialog _____

Questel/Other _____

Or Link _____

Lexis Nexis _____

Sequence Systems 04/05p

WWW Internet _____

Other specify _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:00:27 ; Search time 63 Seconds
(without alignments)
4597.003 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSNGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSDED 1025

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	5311	100.0	1025	AAE10155	Aae10155 Human mel
2	5285	99.5	1025	AAM47798	Aam47798 Human RNA
3	5285	99.5	1025	ADC31794	Adc31794 Human nov
4	4306	81.1	838	AAE10165	Aae10165 RNA helic
5	2595.5	48.9	558	AAU23090	Aau23090 Novel hum
6	2456	46.2	468	ADA54899	Ada54899 Human pro
7	1860	35.0	417	AAU23647	Aau23647 Novel hum
8	1839	34.6	356	AAM40129	Aam40129 Human pol
9	1769	33.3	348	AAM41915	Aam41915 Human pol
10	1769	33.3	348	ADC33339	Adc33339 Human nov
11	1352	25.5	678	ABP69672	Abp69672 Human pol
12	1344	25.3	678	AAU93708	Aau93708 Human pro
13	1093	20.6	304	AAU00296	Aau00296 Interfero
14	1093	20.6	308	AAU00298	Aau00298 Interfero
15	861	16.2	267	AAU00297	Aau00297 Interfero
16	861	16.2	271	AAU00299	Aau00299 Interfero
17	808	15.2	448	AAU23654	Aau23654 Novel hum
18	783	14.7	166	ABB42219	Abb42219 Peptide #
19	783	14.7	166	AAM36026	Aam36026 Peptide #
20	783	14.7	166	AAM75915	Aam75915 Human bon
21	783	14.7	166	AAM63104	Aam63104 Human bra
22	782	14.7	154	AAG74314	Aag74314 Human col
23	782	14.7	154	ABP41112	Abp41112 Human ova
24	666.5	12.5	357	AAH42981	Aah42981 Human ORF
25	563.5	10.6	447	ABP41171	Abp41171 Human ova

26	558.5	10.5	447	4	AAU95485	Aab95485 Human pro
27	521	9.8	312	4	AAU23099	Aau23099 Novel hum
28	488	9.2	92	4	AAM14138	Aam14138 Peptide #
29	488	9.2	92	4	ABB33083	Abb33083 Peptide #
30	488	9.2	92	4	AAM26544	Aam26544 Peptide #
31	488	9.2	92	4	ABB27911	Abb27911 Human pep
32	488	9.2	92	4	ABB18551	Abb18551 Protein #
33	488	9.2	92	4	AAM66268	Aam66268 Human bon
34	488	9.2	92	4	AAM53880	Aam53880 Human bra
35	488	9.2	92	4	ABG47932	Abg47932 Human liv
36	488	9.2	92	4	AAM01876	Aam01876 Peptide #
37	488	9.2	92	5	ABG35915	Abg35915 Human pep
38	453	8.5	752	4	AAU96107	Aau96107 Putative
39	406	7.6	1909	4	AAU31166	Aab31166 Amino aci
40	395.5	7.4	764	4	AAU62028	Aab62028 Recombina
41	357	6.7	71	4	ABB42510	Abb42510 Peptide #
42	357	6.7	71	4	AAM36323	Aam36323 Peptide #
43	357	6.7	71	4	AAM76213	Aam76213 Human bon
44	332.5	6.3	1383	4	ABB63586	Abb63586 Drosophil
45	332	6.3	1924	4	AAU09768	Aae09768 Human dic

ALIGNMENTS

RESULT 1
AAE10155
ID AAE10155 standard; protein; 1025 AA.
XX
AC AAE10155;
XX
DT 29-NOV-2001 (first entry)
XX
DE Human melanoma differentiation associated (Mda)-5 protein.
XX
KW Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;
KW RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
KW central nervous system; cytostatic; apoptosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 125..174
FT /note= "Reveals sequence homology with other CARD
FT proteins"
FT Domain 331..336
FT /note= "ATPase A motif"
FT Domain 443..446
FT /note= "ATPase B motif"
FT Region 722..823
FT /note= "Reveals significant homology to RNA helicase C-
XX terminal conserved domain"
PN WO200164707-A1.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US0006960.
XX
PR 29-FEB-2000; 2000US-00515363.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Fisher PB, Kang D, Gopalkrishnan RV;
XX
DR WPI; 2001-565494/63.
DR N-PSDB; AAD17203.
XX
PT Nucleic acid sequences encoding a Melanoma Differentiation Associated
PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral
PT activity.

XX Claim 21; Page 18-19; 152pp; English.

PS The present invention relates to an isolated nucleic acid encoding a

XX melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5

CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.

CC Mda-5 is a novel interferon (IFN) inducible gene with structural

CC similarities to RNA helicases and CARD motif containing proteins. Mda-5

CC is induced during terminal differentiation in human melanoma cells

CC treated with the combination of recombinant fibroblast IFN and the

CC antileukaemic compound mezerein (MEZ). Mda-5 is useful for identifying

CC compounds that may induce its expression. Mda-5 is useful for treating

CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma

CC multiforme, cervical cancer, breast cancer, colon cancer, prostate

CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a

CC cancer of the central nervous system and apoptosis. The Mda-5 promoter

CC exhibits melanocyte tissue specificity and minimises systemic toxicity.

CC The present sequence is human Mda-5 protein

XX SQ Sequence 1025 AA;

Query Match 100.0%; Score 5311; DB 4; Length 1025;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCFRVRVMYIQVEPVDLYTLFPAEVEKEIQRTVATSGNQAVE 60

Db 1 MSGYSTDENFRYLISCFRVRVMYIQVEPVDLYTLFPAEVEKEIQRTVATSGNQAVE 60

QY 61 LLLSTLEKGVHGLGWTRFVEALRRTGSPLAARYMNPDLTLPSPFENAHDEYLQLNL 120

Db 61 LLLSTLEKGVHGLGWTRFVEALRRTGSPLAARYMNPDLTLPSPFENAHDEYLQLNL 120

QY 121 LQPTLVKLLVRDVLDCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS 180

Db 121 LQPTLVKLLVRDVLDCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS 180

QY 181 FLNVLRTGNELVQELTSGDCSESAEIEIENLSQVDPQVEEQLLSTVQPNLEKEVWGM 240

Db 181 FLNVLRTGNELVQELTSGDCSESAEIEIENLSQVDPQVEEQLLSTVQPNLEKEVWGM 240

QY 241 ENNSSESSFADSSVSVESDTSLAEGSVCLDESIGHNSNMGSDSGTNGSDSDEENVAARA 300

Db 241 ENNSSESSFADSSVSVESDTSLAEGSVCLDESIGHNSNMGSDSGTNGSDSDEENVAARA 300

QY 301 SEPELQLRPYQMEVAQPALEGKNIICLPTGSKTRVAVYIAKHLDKKKASEPGKVI 360

Db 301 SEPELQLRPYQMEVAQPALEGKNIICLPTGSKTRVAVYIAKHLDKKKASEPGKVI 360

QY 361 VLNVKLVLEQLFRKEFQPLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420

Db 361 VLNVKLVLEQLFRKEFQPLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420

QY 421 LNLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKENKPVIP 480

Db 421 LNLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKENKPVIP 480

QY 481 LPQILGLTASPGVGGATQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540

Db 481 LPQILGLTASPGVGGATQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540

QY 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAAGKGNRKERV 600

Db 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAAGKGNRKERV 600

QY 601 AEHLRYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDEYCDGDEDE 660

Db 601 AEHLRYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDEYCDGDEDE 660

QY 661 DDLKKPLKLDLTDRLMTLFFENNMKMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720

Db 661 DDLKKPLKLDLTDRLMTLFFENNMKMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720

QY 721 GIIFTKTQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKERT 780

Db 721 GIIFTKTQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKERT 780

QY 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYLVVAHSGSGV 840

Db 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYLVVAHSGSGV 840

QY 841 IEHETVNDFREKMYKAIHCVQNMPKEEYAHKILELQMQSIMKMKTKRNIKHYKNP 900

Db 841 IEHETVNDFREKMYKAIHCVQNMPKEEYAHKILELQMQSIMKMKTKRNIKHYKNP 900

QY 901 SLITFLCKNCSVLACSGEDIHVIKMHVNMTPFEFKELYIVRENKALQKKCADIQINGEI 960

Db 901 SLITFLCKNCSVLACSGEDIHVIKMHVNMTPFEFKELYIVRENKALQKKCADIQINGEI 960

QY 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCL 1020

Db 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCL 1020

QY 1021 FSDDED 1025

Db 1021 FSDDED 1025

RESULT 2

AAM47798

ID AAM47798 standard; protein; 1025 AA.

XX AAM47798;

XX 05-MAR-2002 (first entry)

XX Human RNA helicase RH116.

DE Human; RH116; RNA helicase; cytostatic; virucide; anti-HIV;

KW immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;

KW antiarteriosclerotic; osteopathic; antidiabetic; hepatotropic;

KW antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease;

KW autoimmune disease; graft rejection; vaccine.

XX

OS Homo sapiens.

XX WO200185955-A1.

CC The present sequence is the protein sequence for human RH116. RH116 is a

CC 116kDa protein and has homology to RNA helicases (DEXH box). RH116 and

CC its coding sequence are useful for treating cancer; acute or chronic

CC infections (especially by HIV or hepatitis B or C); inherited genetic

CC diseases; (auto)immune diseases (particularly rheumatism, arthritis,

CC arteriosclerosis, osteoporosis and diabetes, but many others listed) and

CC to prevent graft rejection. RH116 and its coding sequence are also useful

CC for inducing, or increasing, the immune response to a vaccine

CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1025 AA;

Query Match 99.5%; Score 5285; DB 7; Length 1025;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTFLPAEVKEQIQRTVATSGNQAVE 60
Db 1 MSGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTFLPAEVKEQIQRTVATSGNQAVE 60
QY 61 LLLSTLEKGVVHLGWTRFEFVALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL 120
Db 61 LLLSTLEKGVVHLGWTRFEFVALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL 120
QY 121 LQPTLVVDKLLVRDVLDKMEEELLTIEDNRRIAAENNGNESGVRELLKRIVQKENWFS 180
Db 121 LQPTLVVDKLLVRDVLDKMEEELLTIEDNRRIAAENNGNESGVRELLKRIVQKENWFS 180
QY 181 FLNVLRTGNNELVQELTSGDCSESNAELENLSQVDPQVVEQLLSTTVQPNLEKEVWGM 240
Db 181 FLNVLRTGNNELVQELTSGDCSESNAELENLSQVDPQVVEQLLSTTVQPNLEKEVWGM 240
QY 241 ENNSSESSFADSSVSVESDTSLAEGSVCLDESIGHNSNMGSDSGTMGSDSEENVAARA 300
Db 241 ENNSSESSFADSSVSVESDTSLAEGSVCLDESIGHNSNMGSDSGTMGSDSEENVAARA 300
QY 301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDLKXKASEPGKVI 360
Db 301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDLKXKASEPGKVI 360
QY 361 VLNVKVLVLEQLFRKEFQPFLLKMYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420
Db 361 VLNVKVLVLEQLFRKEFQPFLLKMYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420
QY 421 LLNLENGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKNKPVIP 480
Db 421 LLNLENGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKNKPVIP 480
QY 481 LPQLGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
Db 481 LPQLGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
QY 541 IADATREDPFKEKLEIMTRIQTQYCMSPMSDFGTQPYEQWAIQMEKKAAGKGNRKERV 600
Db 541 IADATREDPFKEKLEIMTRIQTQYCMSPMSDFGTQPYEQWAIQMEKKAAGKGNRKERV 600
QY 601 AEHLRKYNALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDDVYCDGDEDE 660
Db 601 AEHLRKYNALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDDVYCDGDEDE 660
QY 661 DDLKKPLKLDDETDRFLMTLFFENNMKRLAENPEYENKLTCLRNTIMEQYTRTEESAR 720
Db 661 DDLKKPLKLDDETDRFLMTLFFENNMKRLAENPEYENKLTCLRNTIMEQYTRTEESAR 720
QY 721 GIIFTKTRQSAYALSQWITENKFAEVGVKAHHLIGAGHSSEFKPMTQNEKEVISKFRT 780
Db 721 GIIFTKTRQSAYALSQWITENKFAEVGVKAHHLIGAGHSSEFKPMTQNEKEVISKFRT 780
QY 781 GKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840
Db 781 GKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840

QY 841 IEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMQSIMKMKTKRNIKHYKNP 900
Db 841 IERETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMQSIMKMKTKRNIKHYKNP 900
QY 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPFEKELYIVRENKALQKCCADYQINGEI 960
Db 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPFEKELYIVRENKALQKCCADYQINGEI 960
QY 961 ICKGQAWGTMMVHKGLDLPCLKIRNFVVVFKNSTKKQYKWKVELPITFPNLDYSECCL 1020
Db 961 ICKGQAWGTMMVHKGLDLPCLKIRNFVVVFKNSTKKQYKWKVELPITFPNLDYSECCL 1020
QY 1021 FSDED 1025
Db 1021 FSDED 1025
RESULT 4
AAE10165
ID AAE10165 standard; protein; 838 AA.
XX
AC AAE10165;
XX
DT 29-NOV-2001 (first entry)
XX
DE RNA helicase conserved motif of human Mda-5 protein.
XX
KW Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;
KW RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
KW breast; colon; prostate; osteosarcoma; Chondrosarcoma; systemic toxicity;
KW central nervous system; cytostatic; apoptosis.
XX
OS Homo sapiens.
XX
PN WO200164707-A1.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006960.
XX
PR 29-FEB-2000; 2000US-00515363.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Fisher PB, Kang D, Gopalkrishnan RV;
XX
DR WPI; 2001-565494/63.
XX
XX Nucleic acid sequences encoding a Melanoma Differentiation Associated
PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral
PT activity.
XX
PS Example 1; Fig 1D; 152pp; English.
XX
CC The present invention relates to an isolated nucleic acid encoding a
CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5
CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.
CC Mda-5 is a novel interferon (IFN) inducible gene with structural
CC similarities to RNA helicases and CARD motif containing proteins. Mda-5
CC is induced during terminal differentiation in human melanoma cells
CC treated with the combination of recombinant fibroblast IFN and the
CC antileukemic compound mezerein (MEZ). Mda-5 is useful for identifying
CC compounds that may induce its expression. Mda-5 is useful for treating
CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma
CC multiforme, cervical cancer, breast cancer, colon cancer, prostate
CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a
CC cancer of the central nervous system and apoptosis. The Mda-5 promoter
CC exhibits melanocyte tissue specificity and minimises systemic toxicity.
CC The present sequence is RNA helicase conserved motif of human Mda-5
CC protein
XX

SQ	Sequence 838 AA;	
Query Match 81.1%; Score 4306; DB 4; Length 838;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	61 LLLSTLEKGVHGLGWTREFVEALRRGTGSPLAARYNNPELTDLPSPSFENAHDEYQLQLNL 120	
Db	1 LLLSTLEKGVHGLGWTREFVEALRRGTGSPLAARYNNPELTDLPSPSFENAHDEYQLQLNL 60	
QY	121 LQPTLVKLLVRDVLDKCMEEBELLTIEDRNRIAAENNGNESGVRELLKRIVQKENVFSA 180	
Db	61 LQPTLVKLLVRDVLDKCMEEBELLTIEDRNRIAAENNGNESGVRELLKRIVQKENVFSA 120	
QY	181 FLNVLRTGNLQVQLTSGDSCSESAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240	
Db	121 FLNVLRTGNLQVQLTSGDSCSESAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 180	
QY	241 ENNSSSESFADSSVSVSESDTSLAEGSVSCDLSLGHNSNMGSDSGTMSDSDSEENVAARA 300	
Db	181 ENNSSSESFADSSVSVSESDTSLAEGSVSCDLSLGHNSNMGSDSGTMSDSDSEENVAARA 240	
QY	301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDLHDKKKKASEPGKVI 360	
Db	241 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDLHDKKKKASEPGKVI 300	
QY	361 VLNVKVLVLEQLFRKEFPFLKKWYRVIGLSGTQLKISFPEVVKSDIIISTAQILENS 420	
Db	301 VLNVKVLVLEQLFRKEFPFLKKWYRVIGLSGTQLKISFPEVVKSDIIISTAQILENS 360	
QY	421 LLNLENGEDAGVQLSDPSLIIDECHHTNKNAVYNNIMRHYLMQKLNNRLKKNKPVIP 480	
Db	361 LLNLENGEDAGVQLSDPSLIIDECHHTNKNAVYNNIMRHYLMQKLNNRLKKNKPVIP 420	
QY	481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFA 540	
Db	421 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFA 480	
QY	541 IADATREDPFKEKLEIMTRIQTICQMSPMSPDFTQPYEQWAIQMEKKAACKGNRKERV 600	
Db	481 IADATREDPFKEKLEIMTRIQTICQMSPMSPDFTQPYEQWAIQMEKKAACKGNRKERV 540	
QY	601 AEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDDSDEGGDDEYCDGDE 660	
Db	541 AEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDDSDEGGDDEYCDGDE 600	
QY	661 DDLKKPLKLDLTDRLMTLFFENNMKLRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720	
Db	601 DDLKKPLKLDLTDRLMTLFFENNMKLRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 660	
QY	721 GIIFTKTRQSAVALSOWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFT 780	
Db	661 GIIFTKTRQSAVALSOWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFT 720	
QY	781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840	
Db	721 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 780	
QY	841 IEHETVNDFREKMMYKAHCVQNMKPPEYAHKILELQMQSIMCKMKTNRNIAKHYN 898	
Db	781 IEHETVNDFREKMMYKAHCVQNMKPPEYAHKILELQMQSIMCKMKTNRNIAKHYN 838	
RESULT 5		
AAU23090		
ID	AAU23090 standard; protein; 558 AA.	
XX		
AC	AAU23090;	
XX		
DT	18-DEC-2001 (first entry)	
XX		
DE	Novel human enzyme polypeptide #176.	
XX		

KW	Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW	ligase; hyperproliferative disorder; immunodeficiency disorder;
KW	autoimmune disorder; neurological disorder; metabolic disorder;
KW	inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW	blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW	nephrotropic; anticoagulant.
XX	
OS	Homo sapiens.
XX	WO200155301-A2.
PN	
XX	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US001239.
XX	
PR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.
PR	19-MAY-2000; 2000US-0205515P.
PR	07-JUN-2000; 2000US-0209467P.
PR	28-JUN-2000; 2000US-0214886P.
PR	30-JUN-2000; 2000US-0215135P.
PR	07-JUL-2000; 2000US-0216647P.
PR	07-JUL-2000; 2000US-0216880P.
PR	11-JUL-2000; 2000US-0217487P.
PR	11-JUL-2000; 2000US-0217496P.
PR	14-JUL-2000; 2000US-0218290P.
PR	26-JUL-2000; 2000US-0220963P.
PR	26-JUL-2000; 2000US-0220964P.
PR	14-AUG-2000; 2000US-0224518P.
PR	14-AUG-2000; 2000US-0224519P.
PR	14-AUG-2000; 2000US-0225213P.
PR	14-AUG-2000; 2000US-0225214P.
PR	14-AUG-2000; 2000US-0225266P.
PR	14-AUG-2000; 2000US-0225267P.
PR	14-AUG-2000; 2000US-0225268P.
PR	14-AUG-2000; 2000US-0225270P.
PR	14-AUG-2000; 2000US-0225447P.
PR	14-AUG-2000; 2000US-0225757P.
PR	14-AUG-2000; 2000US-0225758P.
PR	14-AUG-2000; 2000US-0225759P.
PR	18-AUG-2000; 2000US-0226279P.
PR	22-AUG-2000; 2000US-0226681P.
PR	22-AUG-2000; 2000US-0226868P.
PR	22-AUG-2000; 2000US-0227182P.
PR	23-AUG-2000; 2000US-0227009P.
PR	30-AUG-2000; 2000US-0228924P.
PR	01-SEP-2000; 2000US-0229287P.
PR	01-SEP-2000; 2000US-0229343P.
PR	01-SEP-2000; 2000US-0229344P.
PR	01-SEP-2000; 2000US-0229345P.
PR	05-SEP-2000; 2000US-0229509P.
PR	05-SEP-2000; 2000US-0229513P.
PR	06-SEP-2000; 2000US-0230437P.
PR	06-SEP-2000; 2000US-0230438P.
PR	08-SEP-2000; 2000US-0231242P.
PR	08-SEP-2000; 2000US-0231243P.
PR	08-SEP-2000; 2000US-0231244P.
PR	08-SEP-2000; 2000US-0231413P.
PR	08-SEP-2000; 2000US-0231414P.
PR	08-SEP-2000; 2000US-0232080P.
PR	08-SEP-2000; 2000US-0232081P.
PR	12-SEP-2000; 2000US-0231968P.
PR	14-SEP-2000; 2000US-0232397P.
PR	14-SEP-2000; 2000US-0232398P.
PR	14-SEP-2000; 2000US-0232399P.
PR	14-SEP-2000; 2000US-0232400P.
PR	14-SEP-2000; 2000US-0232401P.
PR	14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465566/50.
-R N-PSDB; AAS40960.
XX
PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
PS Claim 11; SEQ ID NO 1086; 1180pp; English.
XX
CC The present invention relates to the isolation of novel human enzyme
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. Influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAU22915-AAU23814 represent
CC the novel human enzyme polypeptides of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 558 AA;

Query Match 48.9%; Score 2595.5; DB 4; Length 558;
Best Local Similarity 94.6%; Pred. No. 5.7e-192;
Matches 510; Conservative 10; Mismatches 16; Indels 3; Gaps 2;

QY 294 ENVAARASPEPELQRLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKA 353
Db 1 ENVAARASPEPELQRLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKA 60

QY 354 SEPGKVVIVLVNKVLLVBQLFRKEFPFLKQVYRVLGSGDTQLKISFPEVVKSCDIIIST 413
Db 61 SEPGKVVIVLVNKVLLVBQLFRKEFPFLKQVYRVLGSGDTQLKISFPEVVKSCDIIIST 120

QY 414 AQILENSLLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKK 473
Db 121 AQILENSLLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKK 180

QY 474 ENKPVIPLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQ 533
Db 181 ENKPVIPLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQ 240

QY 534 EPCKKFAIADATREDPFKEKLLEIMTRIQYQMSDPMDFGTQPYEQWAIQMEKKAACKG 593
Db 241 EPCKKFAIADATREDPFKEKLLEIMTRIQYQMSDPMDFGTQPYEQWAIQMEKKAACKG 300

QY 594 NRKERVCAEHLRKYNEALQINDTIRMIDAYTHLETTFYNEEKDKKFAVIEDSDDEGGDDEY 653
Db 301 NRKERVCAEHLRKYNEALQINDTIRMIDAYTHLETTFYNEEKDKKFAVIEDSDDEGGDDEY 360

QY 654 CDGDEDEDLKKPLKLDLTDRLMTLFFENNKMLKRLAENPEYENEKLTCLRNTIMEQYT 713
Db 361 CDGDEDEDLKKPLKLDLTDRLMTLFFENNKMLKRLAENPEYENEKLTCLRNTIMEQYT 420

QY 714 RTEESARGIIFTKTRQSAVALSQWITENKFAEVGVKAHHLIGAGHSSEFKPMTQNEQE 773

Db 421 RTEESARGIIFTKTRQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKE 480
QY 774 VISKFRGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAM-VQARGRARADESTYV 831
Db 481 VISKFRGKINLLIAPQWQKKVILK--NVTLSVMVSSPMKYHGPARGRARADESTYV 537
RESULT 6
ADA54899
ID ADA54899 standard; protein; 468 AA.
XX
AC ADA54899;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 2467.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-395539/38.
DR N-PSDB; ADA53260.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 2467; 205pp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 468 AA;
Query Match 46.2%; Score 2456; DB 6; Length 468;
Best Local Similarity 99.1%; Pred. No. 2.8e-181;
Matches 464; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 558 MTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKERVCAEHLRKYNEALQINDTI 617
Db 1 MTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKERVCAEHLRKYNEALQINDTI 60
QY 618 RMDIAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDDLLKKPLKLDETDRFLM 677
Db 61 RMDIAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDDLLKKPLKLDETDRFLM 120
QY 678 TLFFENNMKRLAENPEYENKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAVALSQW 737
Db 121 TLFFENNMKRLAENPEYENKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAVALSQW 180

QY 738 ITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGKINLLIATTVAEEGLD 797
Db 181 ITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGKINLLIATTVAEEGLD 240
QY 798 IKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGVIEHETVNDFREKMMYKA 857
Db 241 IKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGVIERETVNDFREKMMYKA 300
QY 858 IHCQNMKPEEYAHKILBLQMSIMEKKMKTNRNIKAKHYKNNPSLITFLCKNCVSLACSG 917
Db 301 IHCQNMKPEEYAHKILBLQMSIMEKKMKTNRNIKAKHYKNNPSLITFLCKNCVSLACSG 360
QY 918 EDIHVIEKMHVNMTPPEFKELYIVRENKALQKKCADYQINGEIIICKCGQAWGTMVHKGL 977
Db 361 EDIHVIEKMHVNMTPPEFKELYIVRENKTLQKKCADYQINGEIIICKCGQAWGTMVHKGL 420
QY 978 DLPCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCLFSD 1025
Db 421 DLPCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCLFSD 468
RESULT 7
AAU23647
ID AAU23647 standard; protein; 417 AA.
XX
AC AAU23647;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human enzyme polypeptide #733.
XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW nephrotropic; anticoagulant.
XX
OS Homo sapiens.
XX
PN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001239.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465566/50.
N-PSDB; AAS41517.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases.

Claim 11; SEQ ID NO 1643; 1180pp; English.

The present invention relates to the isolation of novel human enzyme polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. Influenza). The polynucleotides of the invention can also be used in gene therapy. AAU2915-AAU23814 represent the novel human enzyme polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 417 AA;

Query Match 35.0%; Score 1860; DB 4; Length 417;
Best Local Similarity 90.8%; Pred. No. 3.6e-135;

PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-004711275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI61071.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Example 2; SEQ ID NO 6846; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 348 AA;

Query Match 33.3%; Score 1769; DB 4; Length 348;
Best Local Similarity 97.4%; Pred. No. 3.1e-128;
Matches 337; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 680 FFENNMKRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTQSAVALSQWIT 739
Db 3 FLKTIKXNLRLAHPXYENEKLTCLRNTIMEQYTRTEESARGIIFTKTQSAVALSQWIT 62

QY 740 ENKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGKINLLIATTVAEGLDIK 799
Db 63 ENKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGKINLLIATTVAEGLDIK 122

QY 800 ECNIVIRYGLVTNEIAMVQARGRADESTYVLVAHSGSGVIEHETVNDFREKMYKAH 859
Db 123 ECNIVIRYGLVTNEIAMVQARGRADESTYVLVAHSGSGVIEHETVNDFREKMYKAH 182

QY 860 CVQNMKPEEYAHKILELQMQSIMKKMKTNRNIAKHYKNPISLITFLCKNCVSLACSGED 919
Db 183 CVQNMKPEEYAHKILELQMQSIMKKMKTNRNIAKHYKNPISLITFLCKNCVSLACSGED 242

QY 920 IHVIEKMHVNMTPFEKELYIVRENKALQKKADYQINGEITCKGQAWGTMMVHKGLDL 979
Db 243 IHVIEKMHVNMTPFEKELYIVRENKTLQKKADYQINGEITCKGQAWGTMMVHKGLDL 302

QY 980 PCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCLFSDED 1025

Db 303 PCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCLFSDED 348
|||||
RESULT 10
ADC33339
ID ADC33339 standard; protein; 348 AA.
XX
AC ADC33339;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:3421.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR N-PSDB; ADC32572.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Example 2; SEQ ID NO 3421; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

Db 494 LINEALETLMQAAVAQKMDQAEYQAKIRDLQQAALTKRAAQAAQRENQRQFPVEHVQ 553
QY 905 FLCNCSVLACSGEDIHVIEKMHVNMTPFEKELY-IVRENKALOKKCADYQINGEILICK 963
Db 554 LLCINCMVAVGHGSDLRKVEGTHVNVNPNFNSYNNVSRDPVIVNKVFKDWKPGGVISCR 613
QY 964 -CGQAWGTMVHKGLDLPCLKIRNFVVFKNNSTKKQYKKWVELPITFPNLDYSECC 1019
Db 614 NCGEVWGLQMIYKSVKLPVLKVR--MLLETPOGRIQAKKWSRVPFVSPDFDFLQHC 668

RESULT 12
AAB93708
ID AAB93708 standard; protein; 678 AA.
XX
AC AAB93708;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:13299.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
Claim 8; SEQ ID NO 13299; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 678 AA;
Query Match 25.3%; Score 1344; DB 4; Length 678;
Best Local Similarity 40.6%; Pred. No. 7.5e-95;
Matches 291; Conservative 136; Mismatches 238; Indels 52; Gaps 12;
QY 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHDKKASEPGKVVIVLNK 365
Db 1 MELRSYQWEVIMPALGKNIICLPTGAGKTRAAAYVAKRHL-----TVDGAKVVVIVNR 56
QY 366 VLLVEQLFRKEFQFELKKVYRVLGSDTQLKISFPEVVKSCDIIISTAIENSLNLE 425
Db 57 VHLVTQ-HGEFFRMLDGRWTVTTLSGDMGPRAGFGLARCHDLLICTAELLQWALTSP 115
QY 426 NGEDAGVQLSDFSLIIIDECHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIPLPQIL 485
Db 116 --EEHVELTVFSLIVDECHTHKDTVYNVIMSQYLELKLQRAQ-----PLPQVL 164
QY 486 GLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFAIADAT 545
Db 165 GLTASPGTGASKLDGAINHVLQLCANLDTWCIMSPQCCPOLQEHSSQPCCKQYNLCHRR 224
QY 546 REDPFKEKLEIMTRIQTYCQMPMS-DFGTQPYEQWAIQMEKAAKGNRKERVCAEHL 604
Db 225 SQDPFGDLLKLMQDIHDLHLEMPELSRKFGTQMYEQVVKLSEAAALAGLQEQRYVALHL 284
QY 605 RKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDEGDDDEYCDGDEDDLK 664
Db 285 RRYNDALLIHDTVRAVDALAAALQDFYHREHVTQTIL-----C----- 322
QY 665 KPLKLDETDRFLMTLFFENKMKRLAENPEYENKLTKLNTIMEQYTRTEESARGIIF 724
Db 323 -----AERRLLALFDDRKNELAHATHGP-ENPKLEMLEKILQRFSS-SSNSPRGIIF 373
QY 725 TKTRQSAVALSQWITENEFKFAEVGVKAHHLIGAGHSSEFPKPMTONQEKVISKFTGKIN 784
Db 374 TRTRQSAHSLLLWLQOQGLQTVDIRAQLLIGAGNSSQSTHMTQDQEQEVIQEQDGTIN 433
QY 785 LLIATTVABEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEHE 844
Db 434 LLVATSAEGLDIPHCNVVVRYGLLTNEISMVQARGAWADQSVYAFVATEGSRRLKRE 493
QY 845 TVNDFREKMYKAIHCVQNMKPEYAHKILELQMQSIMEKMKTKRNIAXHYKNPSLIT 904
Db 494 LINEALETLMQAAVAQKMDQAEYQAKIRDLQQAALTKRAAQAAQRENQRQFPVEHVQ 553
QY 905 FLCNCSVLACSGEDIHVIEKMHVNMTPFEKELY-IVRENKALOKKCADYQINGEILICK 963
Db 554 LLCINCMVAVGHGSDLRKVEGTHVNVNPNFNSYNNVSRDPVIVNKVFKDWKPGGVISCR 613
QY 964 -CGQAWGTMVHKGLDLPCLKIRNFVVVKNNSTKKQYKKWVELPITFPNLDYSECC 1019
Db 614 NCGEVWGLQMIYKSVKLPVLKVR--MLLETPOGRIQAKKWSRVPFVSPDFDFLQHC 668

RESULT 13
AAU00296
ID AAU00296 standard; protein; 304 AA.
XX
AC AAU00296;
XX
DT 12-SEP-2001 (first entry)
XX
DE Interferon induced polypeptide, IFN4.

KW Interferon induced nucleic acid; autoimmune disease; lupus erythematosus; immunodeficiency; stroke; acquired immunodeficiency syndrome; AIDS;
KW graft rejection; viral infection; hepatitis; aplastic anaemia; cancer;
KW human immunodeficiency virus; HIV; immune-mediated glomerulonephritis;
KW haematologic disease; chronic neutropenia; myocardial infarction;
KW neurological disease; Alzheimer's disease; Parkinson's disease; tumour;
KW amyotrophic lateral sclerosis; spinal muscular atrophy; human; IFN4.

XX Homo sapiens.
OS
XX WO200118208-A2.
PN
XX 15-MAR-2001.
PD
XX 08-SEP-2000; 2000WO-US024704.
PF
XX 08-SEP-1999; 99US-0152921P.
PR
XX 20-OCT-1999; 99US-0160575P.
PR
XX 20-JAN-2000; 2000US-0177104P.
PR
XX 07-SEP-2000; 2000US-00656633.
PR
XX (CURA-) CURAGEN CORP.
PA
XX (BIOJ) BIOGEN INC.
PA
XX Peyton JA, Da Silva A, Hochman P, Hsu A;
PI
XX WPI; 2001-235201/24.
DR
XX N-PSDB; AAS01149.
DR
XX New interferon induced polypeptides and polynucleotides, useful for the
PT diagnosis, prevention and treatment of immunological, cell proliferative
PT disorders, such as lupus erythematosus, cancer, stroke and Alzheimer's
PT disease.
XX Claim 1; Page 29-32; 134pp; English.
PS
XX The sequence represents interferon induced polypeptide, IFN4. IFN nucleic
XX acids and polypeptides are useful for treating or preventing a pathology
CC associated with IFN polypeptide in a human. They are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of IFN polypeptide or polynucleotide. IFN nucleic
CC acids, polypeptides and antibodies are useful for diagnosis, prevention
CC or treatment of variety of immunological and cell proliferative
CC disorders, such as autoimmune diseases e.g. lupus erythematosus,
CC immunodeficiency diseases such as acquired immunodeficiency syndrome
CC (AIDS), graft rejection, viral infections including hepatitis and human
CC immunodeficiency virus (HIV), immune-mediated glomerulonephritis,
CC haematologic diseases such as aplastic anaemia and chronic neutropenia
CC and cancer. In addition they are also useful for treating or diagnosing
CC various disorders associated with cell death, including myocardial
CC infarction, stroke, neurological diseases including Alzheimer's and
CC Parkinson's diseases, amyotrophic lateral sclerosis and spinal muscular
CC atrophy. IFN nucleic acids and polypeptides are also useful for
CC identifying interferon-like proteins and interferon agonists, for
CC screening drugs and compounds which inhibit or enhance IFN activity or
CC function and as targets for the identification of small molecules that
CC are immunostimulatory, immunosuppressive, or stimulate or suppress normal
CC cell or tumour cell growth in mammals, including humans
XX Sequence 304 AA;
SQ
Query Match 20.6%; Score 1093; DB 4; Length 304;
Best Local Similarity 100.0%; Pred. No. 6e-76;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 EVKEIQRTVATSGNQAVELLSTLEKGVHGLGWTREFVEALRRTGSPLAARYMNPFLT 100
Db |
87 EVKEIQRTVATSGNQAVELLSTLEKGVHGLGWTREFVEALRRTGSPLAARYMNPFLT 146
QY 101 DLPSPSFENAHDEYQLNLLQPTLVLDKLVDRDVKCMEEELLTIEDNRNIAAENNGN 160
Db |
147 DLPSPSFENAHDEYQLNLLQPTLVLDKLVDRDVKCMEEELLTIEDNRNIAAENNGN 206
QY 161 ESGVRELLKRIVQKNWFSFAFLNVLRTQGNNELVQELTGDSCSENAEINLSQVDGPQV 220
Db |
207 ESGVRELLKRIVQKNWFSFAFLNVLRTQGNNELVQELTGDSCSENAEINLSQVDGPQV 266
QY 221 EEQLLSTTVQPNLEKEVWGMENNSSESSFADSSVVS 256
Db |
267 EEQLLSTTVQPNLEKEVWGMENNSSESSFADSSVVS 302

RESULT 14
AAU00298
ID AAU00298 standard; protein; 308 AA.
XX
AC AAU00298;
XX
DT 12-SEP-2001 (first entry)
XX
DE Interferon induced polypeptide, IFN6.
XX
KW Interferon induced nucleic acid; autoimmune disease; lupus erythematosus;
KW immunodeficiency; stroke; acquired immunodeficiency syndrome; AIDS;
KW graft rejection; viral infection; hepatitis; aplastic anaemia; cancer;
KW human immunodeficiency virus; HIV; immune-mediated glomerulonephritis;
KW haematologic disease; chronic neutropenia; myocardial infarction;
KW neurological disease; Alzheimer's disease; Parkinson's disease; tumour;
KW amyotrophic lateral sclerosis; spinal muscular atrophy; human; IFN6.
XX
OS Homo sapiens.
XX
PN WO200118208-A2.
XX
PD 15-MAR-2001.
XX
PF 08-SEP-2000; 2000WO-US024704.
XX
PR 08-SEP-1999; 99US-0152921P.
PR 20-OCT-1999; 99US-0160575P.
PR 20-JAN-2000; 2000US-0177104P.
PR 07-SEP-2000; 2000US-00656633.
XX
PA (CURA-) CURAGEN CORP.
PA (BIOJ) BIOGEN INC.
XX
PI Peyton JA, Da Silva A, Hochman P, Hsu A;
XX
DR WPI; 2001-235201/24.
DR N-PSDB; AAS01151.
XX
PT New interferon induced polypeptides and polynucleotides, useful for the
PT diagnosis, prevention and treatment of immunological, cell proliferative
PT disorders, such as lupus erythematosus, cancer, stroke and Alzheimer's
PT disease.
XX Claim 1; Page 40-42; 134pp; English.
PS
XX The sequence represents interferon induced polypeptide, IFN6. IFN nucleic
XX acids and polypeptides are useful for treating or preventing a pathology
CC associated with IFN polypeptide in a human. They are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of IFN polypeptide or polynucleotide. IFN nucleic
CC acids, polypeptides and antibodies are useful for diagnosis, prevention
CC or treatment of variety of immunological and cell proliferative
CC disorders, such as autoimmune diseases e.g. lupus erythematosus,
CC immunodeficiency diseases such as acquired immunodeficiency syndrome
CC (AIDS), graft rejection, viral infections including hepatitis and human
CC immunodeficiency virus (HIV), immune-mediated glomerulonephritis,
CC haematologic diseases such as aplastic anaemia and chronic neutropenia
CC and cancer. In addition they are also useful for treating or diagnosing
CC various disorders associated with cell death, including myocardial
CC infarction, stroke, neurological diseases including Alzheimer's and
CC Parkinson's diseases, amyotrophic lateral sclerosis and spinal muscular
CC atrophy. IFN nucleic acids and polypeptides are also useful for
CC identifying interferon-like proteins and interferon agonists, for
CC screening drugs and compounds which inhibit or enhance IFN activity or
CC function and as targets for the identification of small molecules that
CC are immunostimulatory, immunosuppressive, or stimulate or suppress normal
CC cell or tumour cell growth in mammals, including humans
XX Sequence 308 AA;
SQ

Query Match 20.6%; Score 1093; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 6.1e-76;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EVKEQIORTVATSGNMQAVELLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNP 100
Db 91 EVKEQIORTVATSGNMQAVELLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNP 150

QY 101 DLPSPSFENAHDEYLLQLNLLOPTLVLDKLVDRDVLDCMEEELLTIEDNRIRIAAENNGN 160
Db 151 DLPSPSFENAHDEYLLQLNLLOPTLVLDKLVDRDVLDCMEEELLTIEDNRIRIAAENNGN 210

QY 161 ESGVRELLKRIVQKENWFSAFNLVLRQTGNNELVQELTGSDCSSENAEIEENLSQVDGPQV 220
Db 211 ESGVRELLKRIVQKENWFSAFNLVLRQTGNNELVQELTGSDCSSENAEIEENLSQVDGPQV 270

QY 221 BEQLLSTTVQPNLEKEVWGMENSSSESFADSSVVS 256
Db 271 BEQLLSTTVQPNLEKEVWGMENSSSESFADSSVVS 306

RESULT 15
AAU00297
ID AAU00297 standard; protein; 267 AA.
XX AC AAU00297;
XX DT 12-SEP-2001 (first entry)
XX DE Interferon induced polypeptide, IFN5.
XX KW Interferon induced nucleic acid; autoimmune disease; lupus erythematosus;
KW immunodeficiency; stroke; acquired immunodeficiency syndrome; AIDS;
KW graft rejection; viral infection; hepatitis; aplastic anaemia; cancer;
KW human immunodeficiency virus; HIV; immune-mediated glomerulonephritis;
KW haematologic disease; chronic neutropenia; myocardial infarction;
KW neurological disease; Alzheimer's disease; Parkinson's disease; tumour;
KW amyotrophic lateral sclerosis; spinal muscular atrophy; human; IFN5.
XX OS Homo sapiens.
XX PN WO200118208-A2.
XX PD 15-MAR-2001.
XX PF 08-SEP-2000; 2000WO-US024704.
XX PR 08-SEP-1999; 99US-0152921P.
PR 20-OCT-1999; 99US-0160575P.
PR 20-JAN-2000; 2000US-0177104P.
PR 07-SEP-2000; 2000US-00656633.
XX (CURA-) CURAGEN CORP.
PA (BIOJ) BIOGEN INC.
XX Peyman JA, Da Silva A, Hochman P, Hsu A;
PI WPI; 2001-235201/24.
DR N-PSDB; AAS01150.
XX New interferon induced polypeptides and polynucleotides, useful for the
PT diagnosis, prevention and treatment of immunological, cell proliferative
PT disorders, such as lupus erythematosus, cancer, stroke and Alzheimer's
PT disease.
XX Claim 1; Page 33-35; 134pp; English.
XX The sequence represents interferon induced polypeptide, IFN5. IFN nucleic
CC acids and polypeptides are useful for treating or preventing a pathology
CC associated with IFN polypeptide in a human. They are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of IFN polypeptide or polynucleotide. IFN nucleic
CC acids, polypeptides and antibodies are useful for diagnosis, prevention

CC or treatment of variety of immunological and cell proliferative
CC disorders, such as autoimmune diseases e.g. lupus erythematosus,
CC immunodeficiency diseases such as acquired immunodeficiency syndrome
CC (AIDS), graft rejection, viral infections including hepatitis and human
CC immunodeficiency virus (HIV), immune-mediated glomerulonephritis,
CC haematologic diseases such as aplastic anaemia and chronic neutropenia
CC and cancer. In addition they are also useful for treating or diagnosing
CC various disorders associated with cell death, including myocardial
CC infarction, stroke, neurological diseases including Alzheimer's and
CC Parkinson's diseases, amyotrophic lateral sclerosis and spinal muscular
CC atrophy. IFN nucleic acids and polypeptides are also useful for
CC identifying interferon-like proteins and interferon agonists, for
CC screening drugs and compounds which inhibit or enhance IFN activity or
CC function and as targets for the identification of small molecules that
CC are immunostimulatory, immunosuppressive, or stimulate or suppress normal
CC cell or tumour cell growth in mammals, including humans

XX SQ Sequence 267 AA;
Query Match 16.2%; Score 861; DB 4; Length 267;
Best Local Similarity 96.6%; Pred. No. 4.7e-58;
Matches 170; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 41 EVKEQIORTVATSGNMQAVELLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNP 100
Db 87 EVKEQIORTVATSGNMQAVELLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNP 146

QY 101 DLPSPSFENAHDEYLLQLNLLOPTLVLDKLVDRDVLDCMEEELLTIEDNRIRIAAENNGN 160
Db 147 DLPSPSFENAHDEYLLQLNLLOPTLVLDKLVDRDVLDCMEEELLTIEDNRIRIAAENNGN 206

QY 161 ESGVRELLKRIVQKENWFSAFNLVLRQTGNNELVQELTGSDCSSENAEIEENLSQVD 216
Db 207 ESGVRELLKRIVQKENWFSAFNLVLRQTGNNELVQELTGSDCSSENAEIEENLSQVD 262

Search completed: June 2, 2004, 19:10:14
Job time : 67 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:08:58 ; Search time 23 Seconds
(without alignments)
2300.723 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSNGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSDSD 1025

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	286	5.4	503	US-09-408-020-66	Sequence 66, Appl
2	284	5.3	502	US-09-408-020-34	Sequence 34, Appl
3	217.5	4.1	2662	US-09-595-684B-31	Sequence 31, Appl
4	202	3.8	3878	US-09-914-259-11	Sequence 11, Appl
5	194.5	3.7	2482	US-08-328-254-6	Sequence 6, Appl
6	193.5	3.6	661	US-09-107-532A-3677	Sequence 3677, Ap
7	191.5	3.6	3248	US-08-353-700-1	Sequence 1, Appl
8	191.5	3.6	3248	PCT-US95-16216-1	Sequence 1, Appl
9	188.5	3.5	1388	US-08-685-576-4	Sequence 4, Appl
10	188.5	3.5	1388	US-09-976-594-296	Sequence 296, App
11	187.5	3.5	2125	US-09-919-172-29	Sequence 29, Appl
12	186.5	3.5	1786	US-08-973-462-8	Sequence 8, Appl
13	182.5	3.4	2954	US-09-150-867-1	Sequence 1, Appl
14	179	3.4	666	US-09-134-001C-5465	Sequence 5465, Ap
15	174	3.3	1388	US-08-685-576-1	Sequence 1, Appl
16	173	3.3	1211	US-09-134-001C-4820	Sequence 4820, Ap
17	171.5	3.2	976	US-09-104-324B-4	Sequence 4, Appl
18	171.5	3.2	2285	US-09-308-375-2	Sequence 2, Appl
19	167.5	3.2	1219	US-09-107-532A-6020	Sequence 6020, Ap
20	167	3.1	677	US-09-328-352-4365	Sequence 4365, Ap
21	167	3.1	3696	US-09-134-001C-5080	Sequence 5080, Ap
22	166.5	3.1	1886	US-08-938-105-3	Sequence 3, Appl
23	164	3.1	1354	US-08-685-871-2	Sequence 2, Appl
24	163	3.1	1939	US-09-310-187A-1	Sequence 1, Appl
25	159.5	3.0	956	US-09-134-001C-4452	Sequence 4452, Ap
26	159	3.0	442	US-09-489-039A-11770	Sequence 11770, A
27	158	3.0	781	US-08-675-631-3	Sequence 3, Appl

28	158	3.0	781	4	US-08-359-316A-3	Sequence 3, Appl
29	158	3.0	781	4	US-09-248-776-3	Sequence 3, Appl
30	157.5	3.0	590	4	US-09-543-681A-7931	Sequence 7931, Ap
31	156.5	2.9	1581	4	US-09-866-108A-15754	Sequence 15754, A
32	156	2.9	407	4	US-09-399-081A-11	Sequence 11, Appl
33	156	2.9	961	4	US-09-914-259-66	Sequence 66, Appl
34	155	2.9	454	4	US-09-543-681A-7717	Sequence 7717, Ap
35	155	2.9	1261	3	US-09-208-742-4	Sequence 4, Appl
36	155	2.9	1261	4	US-09-332-295-2	Sequence 2, Appl
37	155	2.9	1261	4	US-09-709-979-2	Sequence 2, Appl
38	155	2.9	1261	4	US-10-147-268-2	Sequence 2, Appl
39	154.5	2.9	875	4	US-09-107-532A-4537	Sequence 4537, Ap
40	154.5	2.9	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
41	153.5	2.9	394	4	US-09-134-000C-4332	Sequence 4332, Ap
42	152	2.9	407	3	US-08-989-370-6	Sequence 6, Appl
43	152	2.9	407	4	US-09-398-169-6	Sequence 6, Appl
44	152	2.9	1111	4	US-09-914-259-28	Sequence 28, Appl
45	151.5	2.9	709	4	US-09-702-705-335	Sequence 335, App

ALIGNMENTS

RESULT 1
US-09-408-020-66
; Sequence 66, Application US/09408020
; Patent No. 6632937
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DORP.002A
; CURRENT APPLICATION NUMBER: US/09/408,020
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-09-408-020-66

Query Match	5.4%	Score 286;	DB 4;	Length 503;
Best Local Similarity	24.7%	Pred. No. 6.6e-16;		
Matches	136;	Conservative 91;	Mismatches 204;	Indels 120; Gaps 24;
QY	303	EP-ELQRPYQMEVAQPALEKNNIIICLPTGSGKTRVAVYIAKHLDKKKASEPGKIV	361	
Db	12	EPGAVERRDYQVGLAEQAIR-ENCIVLPTGLGKTAVALQVISHYLDGRGALFLAPTRV	70	
QY	362	LVNKVLLVEQLFRKEFQPELKKWYRVIGLSGDTQL--KISFPEVVKSC--DIIISTAQIL	417	
Db	71	LVN-----QHRQFLG---RALTISDITLVGTEDTVPRRKKAWGSGVICATPEIT	116	
QY	418	ENSLNLENGEDAGVQLSDPSLIIDECCHTNKEAVYNNIMRHYLMQKLKNNRLKKNKP	477	
Db	117	RNDIAR-----GMVPLEQGLVVFDEHRAVGDYAYSIA-----RAVGENS-	158	
QY	478	VIPLEQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQKNQIQEPCK	537	
Db	159	-----RMIGMTATL-----PSEREKADE-----IMGTLLSKSIAQRTEDD	205	
QY	538	KFATADATREDPFKEKLEIMTRIQTICQMS-PMS-DFGTQPYEQWAIQMEKKAAGKGNRK	596	
Db	206	EWIKVELPPEMKEIQKLK-MALDERYAALKRCGYDLGNSRSLALLRL-RMVVLSGNRR	263	
QY	597	ERVCAEHLRKYNEALQINDTIRMIDAY--THLETIFYNEEKDKKFAVIEDSDSGGDEYC	654	
Db	264	A-----AKPLFTAIRITYALNIFEAGVTPFLKFCERTVKKGAGVAE-----	306	

QY 655 DGEDEDDLKPLKLDLTDRLMTLFFENNKMMLKRLAENPEYENEKLTKLNRNTIMEQYTR 714
Db 307 -----LFEEDR-----NFTGAMARAKAAQAAGMEHPKIPKL----- 337
QY 715 TEESARG-----IIFTKTRQSAVALSQWITENEKFAEYGVKAHHLIG-AGHSSEFKPMTQ 768
Db 338 -BEAVRGAKGKALVFTSYRDSVDLI-----HSKLQAAGINSGLIGKAGE-----KGLKQ 386
QY 769 NEOKEVISKFTGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQAGR-ARADE 827
Db 387 KKQVETVAKFRDGGYDVLVSTRVGEGLDISEVNLVVFYDNPVSSIRYVQRRGRTGRKDA 446
QY 828 STYVLVAHSGS 838
Db 447 GKLVLMAKGT 457
RESULT 2
US-09-408-020-34
; Sequence 34, Application US/09408020
; Patent No. 6632937
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOMP.002A
; CURRENT APPLICATION NUMBER: US/09/408,020
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-09-408-020-34

Query Match 5.3%; Score 284; DB 4; Length 502;
Best Local Similarity 24.3%; Pred. No. 9.8e-16;
Matches 133; Conservative 94; Mismatches 207; Indels 114; Gaps 23;
QY 303 BP-ELQLRPYQMEVAQPALEGKNIICLPTSGKTRVAVYIAKHDLDKXKASEPGKVIV 361
Db 12 BPGAVERRDYQVGLAEQAIR-ENCIVLPTGLKTAVALQVIAHVLDEGRGALFLAPTRV 70
QY 362 LVNKKVLLVEQLFRKEFOPFLKKYRVIGLSGTQL--KISFPEVVKSC--DIIISTAQIL 417
Db 71 LVN-----QHRQFLG---RALTISDITLVGTEDTIPRRKKAWGGSVICATPEIA 116
QY 418 ENSLLNLENGEDAGVQLSDFSLIIDI DECHHTNKAEAVYNNIMRHYLMQKLNNRLKKENKP 477
Db 117 RNDI-----ERGLVPLEQFGLVIFDEAHRAVGVDYAYSSIA-----A 153
QY 478 VIPLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCK 537
Db 154 VGDNSRMVGMTATL-----PSEREKADE---IMGTLLSRSAQRTEDDPDVKPYVQETAT 205
QY 538 KPAIADATREDPFKEKLEIMTRIQTQYQMSMS-DFGTQPYEQWAIQMEKKAAGKGNRK 596
Db 206 EWIKVDLPPEMKEIQRLKLALD-ERYSSLKRCGYDLGNSRSLALLURL-RMVVLGGNRR 263
QY 597 ERVCAEHLRKYNEALQINDTIRMIDAY--THLETIFYNEEKDKKFAVIEDDSDEGGDDEYC 654
Db 264 A-----AKPLFTAIRITYALNIFEAHGVTPFLKFCERTSKKGVGVAE----- 306
QY 655 DGEDEDDLKPLKLDLTDRLMTLFFENNKMMLKRLAENPEYENEKLTKLNRNTIMEQYTR 714
Db 307 -----LFEQDR-----NFTGAIRAKAAQAAGMEHPKIPKLEDAV----- 341
QY 715 TEESARG--IIFTKTRQSAVALSQWITENEKFAEYGVKAHHLIG-AGHSSEFKPMTQNEQ 771

Db 342 --RGARGKALVFTSYRDSVDLI-----HSRLKAAGINSGLIGKAGE-----KGLKQKQ 389
QY 772 KEVISKFTGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQAGR-ARADESTY 830
Db 390 VETVAKFRDGGYDVLVSTRVGEGLDISEVNLVIFYDNPVSSIRYVQRRGRTGRKDA 449
QY 831 VLVAHSGS 838
Db 450 IVLMAKGT 457
RESULT 3
US-09-595-684B-31
; Sequence 31, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 2662
; TYPE: PRT
; ORGANISM: Human
US-09-595-684B-31

Query Match 4.1%; Score 217.5; DB 4; Length 2662;
Best Local Similarity 18.6%; Pred. No. 7.6e-09;
Matches 209; Conservative 185; Mismatches 421; Indels 307; Gaps 47;
QY 9 ENFRYLISCFARVKMYIQVEPVLDTLTFPAEVEKEIQRTVATSGNMQAVELLSTLEK 68
Db 1039 EQQRKIFSLIQEKNELOQMLESVIAEKEQLKTDLKENIEMTIE---NQEELRLGDELKK 1095
QY 69 GVWHLGWTRETV-----EALRRTGSLAARYMNPETLTLPSFSFENAHDEYLOLLNLQ 123
Db 1096 -----QOEIVAEKNHAIKKEGELSRTCDRLAEEVEEKLKESQQLQEQQLLNQVEE 1148
QY 124 TLVDKLLVRDV--LDKCMEEELLTIE--DRNRIAAAEN-NGNESGVRELLK-RIVQKE-- 175
Db 1149 MSEMOKKINEIENLKNELKNKELTLEHETERLELAQKLNENYEEVKSIKERKVKELQ 1208
QY 176 -----NWFSAFLNVLRQTG-----NNELVQELTGSDCSESNAEINL 212
Db 1209 KSFETERDHLRGVIREIEATGLQTKBELKIAHHLKEHETIDELRRS-VSEKTAQIINT 1267
QY 213 SQVDGPQV-----EEQLLSTTVQPNLEKEVWGMENNSSSESFADSSVVSESDTS 261
Db 1268 QDLEKSHTKLQHEIIPVHEEQELL-----PNVKVSETQETMNELELLTEQSTTKDS--TT 1321
QY 262 LAEGSVCLDESIGHNSNMGSDSGTMGSDSDDEENVAARASPEPELQRPYQMEVAQPALE 321
Db 1322 LARIEM---ERLRLNEKF-----QESQEEIKSLTKERDNLKTIKEALEVHKDQL- 1367
QY 322 GKNIIICLPTGSGKTRVAVYIAK-----DHLDKKKKASEPGKVIVLVNKKVLLVEQL 372
Db 1368 -----KEHIRETLAKIQESQSQEQSLNMKEKDNETTKI-----VSEMEQ- 1407
QY 373 FRKEFQPFLLKWRVIRIGLSDGTQLKISFPE---VVKSCDIIISTAQILENSLNLENG-- 427
Db 1408 FPKKDSALLRIEIMLGLS--KRLQESHDEMKSVAKEKDDLQRLQEVQESDQLKENIK 1465

QY 428 EDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIPLPQILGL 487
Db 1466 EIVAKHLEETEELKVAHCCLKEEETIN-----ELRVNLSEKETE----- 1505
QY 488 TASPVGVGATKQAKA-----EEHILKLCANLDAFTIKTVKENLDQKNQIQIEPCCKF 539
Db 1506 -----ISTIQQLEAINDKLNKIQIEIYEKEEQNLKIQISEVQENVNELK-QFKEHRK-- 1557
QY 540 AIADATREDPFKEKLEIMTRIQTQYQ--MSPMSDFGTQPYEQWAIQMEKKAAGKGNR-- 595
Db 1558 --AKDSALQSIESKMLELTNRQLQESQEEIIMIKKEBEMKRVQEQALQIERDQKENTKEI 1615
QY 596 -----KERVCAEHLRKYNEALQIN-----DTIRMID-AYT 624
Db 1616 VAKMKEQKEYQFLKMTAVNETQEKMCIEIHLKEQFETQKLNLENIEFENIRLTQILHE 1675
QY 625 HLETFYNEEKDKKFAVIEDSDDEGDDDEYCDGEDDDLLKPLKLD----- 670
Db 1676 NLEEMRSVTKER-----DOLRSVEETLKVERDQKLENLRETITRDLEKQEBELKIVMHHL 1729
QY 671 -----ETDRFLMTLFFENNKMMLKRLAENPEYENEL-----TKLRNTIMEQYTRTE 716
Db 1730 KEHQETIDKLRGIVSEKTEINSMQKDLHSNDALKAQDLKIQEELRIAHMHLKEQETI 1789
QY 717 ESARGIIFTKTROSAYALSQWITEN-----EKFAEVGVVKAHHLIGAGHS-----SEFK 764
Db 1790 DKLRGIVSEKTDKLSNMQKDLSENSAKLQEKIQELKANEHQILITLKKDVNETQKKVSEME 1849
QY 765 PMTQ--NEQKEVISKFRIGKINLLIATTVAEEGLDIKECNIVIRYGLVNEIAMVQARGR 822
Db 1850 QLKQIKDQSLTSLKLEIENLNL--AQELHENLEEMK-----SVMKERDN 1892
QY 823 ARADESTYVLVAHSGSGVI-----EH-ETVNDPREKMMYKAH 859
Db 1893 LPRVEETIKLERDQKESIQETKARDLEIQEELKTARMLSKHEKTVDKLREKISEKTIQ 1952
QY 860 CVQNMK-----PEEYAHKILELOMQSIM-----EKKMKTKRNIKAHYKNNPSLI 903
Db 1953 ISDIQKDLKSKDELQKIQELQKKEQLLRVKEDVNMSSHKKINEMEQKKQFEPN---- 2008
QY 904 TFLCKNCSVLACSGEDIHVIEKMHVNMVMTPEFKELYIVRENK 945
Db 2009 -YLCK-----CEMDNFQLTQKLHE-----SLEEIRIVAKER 2038

RESULT 4

US-09-914-259-11
; Sequence 11, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 3878
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-914-259-11

Query Match 3.8%; Score 202; DB 4; Length 3878;
Best Local Similarity 18.8%; Pred. No. 3e-07;
Matches 238; Conservative 179; Mismatches 424; Indels 428; Gaps 55;
QY 5 YSTD-----ENFRYLISCFRARKVMIYQVEPVLDTLFLPAEVKEIQRTVATSGNM 56
Db 1136 YSTHVDQVREYMEKDKALCSLKEELIFAQEEKIKELQIHQLEL--QIMKTQETGDEG 1193

QY 57 QAVELLLLSTLEKGVHGLGWTRFVEALRRITGSPLAARYMNPCL-TDLPSPSFENAHDEYL 115
Db 1194 KPLHLLIGLKQKAV-----SEECYFLOTLCVLGEYTYTPALKCEVNAEDKENS GD--- 1244
QY 116 QLLNLLQPTLVD-KLLVVRD-----VLDKCMEE--ELLTIEDR-NRIAAAENNGN--ES 162
Db 1245 YISENEDPELDQYRYEVQDFQENMHTLLNKVTEYNNKLLVLQTRLSKIWQQTDGMKLEF 1304
QY 163 GVRELLKRIVQKENWFSAFNLVLRQTGNNE-----LVQELTSGDCSESNAEIE 211
Db 1305 GEENLPKEETE-----FLSIHSQMTNLEDIDVNHKSKLSLQDLKLEKLEEQVQLE 1357
QY 212 LSQVDGPQVEEQLLSTTVQPNLEKEVWGMENN--SSESSF-----ADSSVVSESDTSLA 263
Db 1358 LIS-----SLQQQLKET--EQNYEAEIHCLQKELQAVSESTVPPSLPVDVSVVITESDAQRT 1411
QY 264 EGSVCLDESIGHNSNMGSDSGTMGSDSDENVAARASPEPELQLRPYQMEVAQPALEGK 323
Db 1412 MYPGSCVKNIDGTIEFSGEGFV---KEETNIV-----KLEKQYQEQLEE---EVA 1457
QY 324 NIIICLP-----TGSGKTRVAVYIAKHLDKKKKASEPGKVIVLVNK 365
Db 1458 KVIVSMSIAFAQQOTELSRISGGKENTASSQAHAHVCOQEHYFNEMKLSQDQ---IGFQT 1514
QY 366 VLLVEQLFRKEFQPLKKWYRVIGLSG-DTQIKISFPEVW---KSC----- 407
Db 1515 FETVDVKFKEEFKPLSKE---LGEHGKEILLNSDPHDIPESKDCVLTISEEMFSKDKT 1570
QY 408 -----DIIISTAQILENSLLNLENGEDAGVQL-----SD 436
Db 1571 FIVRQSIHDEISVSSMDASRLMNEEQLEMRQELVRQYQEHQQAQTORSSIDNENLVSE 1630
QY 437 FSLIIDE-----CHHTNKEAVYNN----- 456
Db 1631 RERVLEELEALKQLSLAGREKLCCELRNSSTQTQNGNENQGEVEEQTFKEKELDRKPED 1690
QY 457 -----IMRHYLMQKLKNNRLKKNKPVIPLPQILGLTASPGVG----- 495
Db 1691 VPPEILSNERYALQK-ANNRLLKILLEVVVKTAAVEETIGRHVLGILDRSSKSQSSASLI 1749
QY 496 -----ATKQAKAEHILKLCANLDAFTIKTVKENLDQKNQIQIEPCCK-----F 539
Db 1750 WRSEAEASVKSCVHEHTRVTDSEIPSYSGSDMPRNDINMWSKVTEGTELSQRLVRS GF 1809
QY 540 AIADATREDPFKEKL-LEIMTRIQTQYQ--MSPMSDFGTQ---PYEQWAIQMEKKA 592
Db 1810 A---GTEIDPENEEMLNMISSRLQAAVEKLELAISETSSQLEHAKVTQELMRSEFRQKQ 1866
QY 593 GNRKERVCAEHLRK--YNE-----ALQINDTIRMIDAYTHLETFFYNEEKDKKFAVI-- 641
Db 1867 EATESLKQOEELRERLHEESRAREQLAVELSKAEGVIDGYADEKTLFERQIQEKTDIIDR 1926
QY 642 -----EDDSDEGGDDE----- 652
Db 1927 LEQELLCASNRLQELEABEQQIQIEERELLSRQKEAMKAEAGPVEQQLQTEKLMKEKLE 1986
QY 653 -YCDGDEDEDDKKPLKLDDET-----RFLMTLFFENNKMMLKRLAENPEYENEKLTCLR 705
Db 1987 VQCAEKVRDDDLQKQVKALEIDVVEEQVSRFT-ELEQEKNTELMDLRQONQALEKQLEKMR 2045
QY 706 NTIMEQYTRTEESARGIIFTKTROSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKP 765
Db 2046 KFLDEQAIDREHERD-----VFQEQIKLEQQLKV-----VPRFQP 2081
QY 766 MTQNEQKEVIS-----KFRGTGINLLIATTVAEEGL--DIKECNIVIRYGLVNEIAMVQ 818
Db 2082 ISEHQTRVEQLANHLKEKTDKCELL---LSKEQLORDIQERN-----EEIEKLE 2129
QY 819 ARGRARADESTYVLVAHSGS-----GVIE----- 842
Db 2130 FRVR-----ELEQALLVSADTFQKVEDRKHFGAVEAKPELSLEVQLQABRDAIDRKEKEIT 2185
QY 843 --HETVNDFREKMMYKAHCVQNMKPEYAHKILELOMQSIMKMKTKR-----NIAKH 895

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 3677:

SEQUENCE CHARACTERISTICS:

LENGTH: 661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...661

SEQUENCE DESCRIPTION: SEQ ID NO: 3677:
US-09-107-532A-3677

Query Match 3.6%; Score 193.5; DB 4; Length 661;
Best Local Similarity 18.8%; Pred. No. 9.4e-08;
Matches 147; Conservative 130; Mismatches 240; Indels 267; Gaps 36;

QY 299 RASPEPEQLRPYQMEVAQP-----ALEGKNIICL-PTGSGKTRVAVYIAKDHL 348
Db 2 RASRHFEL-VSKYPAGDQPEAINQLVDGVGKKAQILLGATGTGKT----YTISNLIE 56
QY 349 KKKKASEPGKVIIVNKKVLLVEQLFR--KEFPFLKKWYRVIGLSGDTQLKISF-----P 401
Db 57 KVNKPT-----LIIAHNKTLAQLYGEFKFFP-----NNAVEYFVSYYDYQP 100
QY 402 EV-VKSCDIIISTAQILENSLLNLENGEDAG-VQLSDFSLLIIDEK-----HHTNKEA 452
Db 101 EAYVPSSDTYIEKSSVNDKLRHSATSSLLERNVDVIVIASVCIFGLGSPFEYQKQV 160
QY 453 V-----YNNIMRHYLMQKLKNN-----RLKKENKVPVLP----- 482
Db 161 VSIRQGAELDRNQLIRDLVSIQFERNDIDFQGRFRVRGVDVVEIFPASRDERALRVEFFG 220
QY 483 -----QILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQL 528
Db 221 DEVERIREVNALTGEVLGETEHVAIFPATHEFTVNDHEHMAVANIK-----ELEQRLTVL 276
QY 529 KNQIQEPCKKFAIADATREDPFKEKLEIMTRIQ-TYQCMSPMSDFG-TQPYEQWAIQME 586
Db 277 RNE-----NKLLEAQLEQRTNYDIEMMLEMGYTSGIENYSRHM 316

QY 587 KKAAGKGNRKERVCAEHLRKYNALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSD 646
Db 317 GR--KEGE-----PPYTLDD-FFPED-----FLIVADESH 343
QY 647 -----EGGDDEYCDGDEDDL-----KKPLKLDLTDTRFLMTLFFENNKM 687
Db 344 VTMPQIRG---MYNGDRARKQMLVDYGFRLPSALDNRPLRLEEFKHV-----NQII 392
QY 688 KRLAENPEYENEK-----LTKLRNTIMEQYTRTEESARGII 723
Db 393 YVSATPGPYEHEQTDTVIQIIRPTGLLDPVIEVRPIMGQIDDLVGEINERVEKQRFV 452
QY 724 FTKTRQAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVSKFRTGKI 783
Db 453 TLTTKMABELTDY-----FKELGIKVKYL-----HSDIKTL---ERTEIIRDLRLGEF 498
QY 784 NLLIATVAEGLDIKECNIVI-----RYGLVTNEIAMVQARGARADESTYVLVAHSGS 838
Db 499 DVLVGINLLREGLDVPVSLVAILDADKEGFLRSERSLVQTIGRAARNEEGKVM----- 553
QY 839 GVIEHETVNDFREKMMYKAHCVQNMKPEEYAHKILELQMSIMEKKMKTKRNIKHYKN 898
Db 554 -----YADKVTD-SMRLAMDETSR-RRTIQQKYNE 581
QY 899 N-----PSLITFLCKNCSVLACSGEDIHVIEKMHVNNMTPEFKELYIVRENKALQ--KKC 951
Db 582 EHGIVPKTIKEIRDLISITKESEDDTKEAVQVSYEENTKEEKDTLLMKLEKMKDAKA 641
QY 952 ADYQ 955
Db 642 LDPE 645

RESULT 7

US-08-353-700-1

; Sequence 1, Application US/08353700

; Patent No. 5599919

; GENERAL INFORMATION:

; APPLICANT: YEN, TIMOTHY J.

; APPLICANT: RATTNER, JEROME B.

; TITLE OF INVENTION: NUCLEIC ACID ENCODING A

; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,

; TITLE OF INVENTION: AND METHODS OF USE

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN

; STREET: 1601 MARKET STREET, SUITE 720

; CITY: PHILADELPHIA

; STATE: PA

; COUNTRY: USA

; ZIP: 19103-2307

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/353,700

; FILING DATE: 09-DEC-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: REED, JANET E.

; REGISTRATION NUMBER: 36,252

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 563-4100

; TELEFAX: (215) 563-4044

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3248 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

Db 1977 DIGDNVAKVNDWKRFLD-----VENELSRIRSEKASIEH-EALYLEADLE 2022

QY 123 PTLVDKL-LVRDVLK-----CMEEELLTI-EDNRRIAAENNGNE--SGVRELLKRIVQ 173

Db 2023 VVQTEKLCLEKNENKQKVIIVCLEBELSVVTSENRQLRGELDTMSKTTALDQLSEKMK 2082

QY 174 K-----ENWFSAFLNVLROT-----GNNELVQELTGTSCSESNAEINLSQ-VDPQVVEEQ 223

Db 2083 KTQELESQSECLHCIOVAEAEVKEKTELLQTLIS-SDVSELLKDKTHLQKELQSLKDSQ 2141

QY 224 LLSTT-----VQPNLEKEVWGMENNS-----SESSFADSSVVSSESDTSLAEGS----- 266

Db 2142 ALSLTKELENQIAQLNKEKELLVKESESLQARLSESDYEKLVNSKALEALVKEGEFAL 2201

QY 267 -VSCLESLGHNSNMGSDSGTMSGSDSE-----ENVAARASPEPELQLRPYQMEVAQP 318

Db 2202 RLSSTQEEV-HQLRRGIEKLRVRIEAEKKQLHIAEKLKERERENDSLKDKVENLERELQ 2260

QY 319 ALEGKNIILCLPTGSGKTRVAVYIAK-DHLDKKKKASEPGKVIVLVNKVLLVEQLFRKEF 377

Db 2261 MSENQELVILDAENSKAEVETLKTQIEEMARSLKIFELDLVTLRSEKENLTKQIQEKQG 2320

QY 378 QPFLKKWVRVIG-----LSGDTQLKISFPPEVVKSCDIIISTAQILENSLLNLEN----- 426

Db 2321 Q--LSELDKLLSSFKSLLEKEQEAIRIQKEESK-----TAVEMLQNLKELNEAVALC 2372

QY 427 GEDAGVQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKENKPVIPLOI-- 484

Db 2373 GDQEIMKATEQSLDPPPIEBEHQLRNSI-----EKLRLARLEADEKKQLCVLOQLKE 2422

QY 485 -----LGLTASPGVGGATKQAKAEHILKL-----CANLD 514

Db 2423 SEHHADLLKGRVENLERELEIARTNQEHAALEAENSKGEVETLKAKIEGMTQSLRGLELD 2482

QY 515 AFTIKTVKENLDQLKNQIQEPCKKFAIADATREDPFKEKLEIMTRIQTICQMSPMDFG 574

Db 2483 VVTIRSEKENLTNELQKEQERISELEIINSSFENILQEK----- 2521

QY 575 TQPYEQWAIQMEKKA-----KKGNRKERVCAEH-----LRKYNEALQIN-- 614

Db 2522 ----EQEKVQMEKSSSTAMEMLOTLQKELN--ERVAALHNDQEAQKAKEQNLSSQVECLE 2575

QY 615 -DTIRMIDAYTHLETFYNEEKDKKFAVIEDDS-----EGGDDEYC--DGDEDEDDLKK 665

Db 2576 LEKAQLLQGLDEAKNYIVLVQSSVKGLIQEVEDGKQKLEKDEEISRLKNQIQOQEQLV 2635

QY 666 PLKLDETTRFLMTLFFENNKMRLKLAENPEY-----ENEKLTCLRNTIMEQYTRTEES 718

Db 2636 KLSQVEGEH--QLWKEQNLELRNLTVLELQKIQVLSQKNASLODTLEVLSQSYKNLENE 2692

QY 719 ARGIIFTKTRQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKF 778

Db 2693 ---LELTRKMDKMSFV-----EKVNKMTAKETELQREMHMAQKTAELQEELSGEKNR 2741

QY 779 RTGKINLL-----IATTVABEGLDIKECNIVIRYGLVTNEIAMVQARGRA 823

Db 2742 LAGELQLLLEBIKSSKDQLKELTLENSELKSLDCMHKDOQVEKGVREEIAEYQLRLHE 2801

QY 824 RADESTYVLVAHSGSGVIEHETVNDFREKMMYKAHCVQNMKPE-----EVAHKILE 875

Db 2802 AEKKHQALLLDINKQYEVEIQT---YREKLTSE-ECLSSQKLEIDLKSSKEELNNSLK 2857

QY 876 LQMOSIMEKKMKT-----RNIAKHVKNPPLITFLCKNSVLACSGEDIHVIEK--- 925

Db 2858 ATTQ-ILEELKTKMDNLYVNVQNLKKNENARAQGMKLLIKSKQCLE---EKEKILQKELS 2913

QY 926 -----MHHVNMTPFEKELYIVRENKA-----LQKKCADYQINGELICK 963

Db 2914 QLQAAQEKQKTGTGMTKVDLTTTEIKELKETLEETKEADEYLDKYCS-LLISHEKLEK 2972

QY 964 CGQAWGTMVH 974

Db 2973 AKEMLETQVAH 2983

RESULT 9

US-08-685-576-4

; Sequence 4, Application US/08685576

; Patent No. 5906819

; GENERAL INFORMATION:

; APPLICANT: Kaibuchi, Koza

; APPLICANT: Iwamatsu, Akihiro

; APPLICANT: Nakano, Takeshi

; APPLICANT: Ito, Masaaki

; APPLICANT: Takahashi, No. 5906819uaki

; TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/685,576

; FILING DATE: 24-JUL-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 7-325129

; FILING DATE: 20-NOV-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 8-17150

; FILING DATE: 05-JAN-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 8-131206

; FILING DATE: 26-APR-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Bent, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 16887/843

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)672-5300

; TELEFAX: (202)672-5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1388 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-685-576-4

Query Match 3.5%; Score 188.5; DB 2; Length 1388;

Best Local Similarity 18.2%; Pred. No. 8.3e-07;

Matches 204; Conservative 195; Mismatches 399; Indels 321; Gaps 57;

QY 71 WHLGWTRFVEALRRRTGSPLAARYMNPDL-TDLPSPSTENAHDEYLLQNLNLLQPTLVVDKL 129

Db 359 WH--W-----DNIRETAAPVV-----PELSSDIDSSNFDIEDD---KGDVETFPPIKA 402

QY 130 LVRDVL-----KCMEEELLTTEDNRRIAAENNGSGVRELLKRIVQKENWFSALNVL 185

Db 403 FVGNQLPFIQFTYYRENLLSDSP---SCRENDISIQRKNEESQEIQKK-----LYTL 452

QY 186 RQTGNNEL-VQELTGSDCSESNAEINLSQVDGPQVEQLLSTTVQPNLEKEVWGMENNS 244

Db 453 EEHLSNEMQAKEELEQCKSVNTRLEKTAK---ELEEEL---TLRKSVESALRQLEK 505

QY 245 SESSFADSSVVSSESDTSL-----AEGSVSCLDLSTL-----GHNSNMGSDS-GTMSGSDS 292


```

506 ALLQHNAEYQRKADHEADKCRNLNLENDVNSLKQLEDLKKRNQNSQISTEKNVQLQRQLD 565
QY 293 EENVAARASPEPELQRLPYQMEVAQPA-----LEGKNIICLPTGSGKTRVAVYIA 343
Db 566 ETNALLRTESDTAARLRKTKQAESSKQIQQLESNNRDLQDKNCL--LETAKLLEKEFINL 623
QY 344 KDHLDKKKKASEPGKVIKLVNKKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISPPEV 403
Db 624 QSALESERRDRTHGSEII-----NDLQG-----RICGLEED----- 654
QY 404 VKSCDIIISTAQI----LENSLLNLENGEDAGVQLS-DFSLIII-----DECHH----- 447
Db 655 LKNGKILLAKVELEKRLQERFTDLEK-EKSNMEIDMTYQLKVIQQSLEQEEAEHKATKA 713
QY 448 --TNKEAVYNNIMRHYLMQKLKNNRLKKENKPVIPLPQILGLTASPGVGGATYQAKAEEH 505
Db 714 RLADKNKIYESI-----EEAKSEAMKEMEKILLE-----ERTLKQKVENL 753
QY 506 IL---KLCANLDAFTIKTVKENLDQLKNQIQEPCKKFAIADATREDPFKEKLLEIMTRIQ 562
Db 754 LLEAEKCSLLDC-----DLKQSQKINELLKQ--KDVLNED-VRNLTLEKEQETQ 801
QY 563 TYCQSPMSDFGTPYEQWAIQMEKKAACKGNRKEVCAEHLRKYNEAL-----QI 613
Db 802 KRC--LTQNDLKMQTQQVNTLKMSEKQLKQENHLMEMKMNLEKQNAELRKEKQDADGQM 859
QY 614 NDTIRMIDAYTHLETFYN-----EEKDK-----KFAVIEDDS- 646
Db 860 KELQDQLEAEQYFSTLYKTVRELKEECBEKTKLGKELQOKKQELQDERDLSLAQLEITL 919
QY 647 -----EGDDEYCDGDED-----EDDLKKPL-----KLDETRFLMTLFFENNMKL 687
Db 920 TKADSEQLARSIAEEQYSDLEKEKIMKELEIKEMMARHKQELTEKDATIASL-EETNRTL 978
QY 688 KRLAENPEYENEKLTCLRNTIMEQYTRTEE--SARGIIFTKTRQSAVALSQWITEN--- 741
Db 979 TSDVANLANEKEELNNKLKDVQEQLSRLKDEEISAAI-----KAQFEKQLLTERTLK 1031
QY 742 ----EKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFTGKINLL----- 786
Db 1032 TQAVNKLAEI-MNRKEPVKRGNDTDVR-RKEKENRKLHMLKSEREKLTOQMIKYQKELN 1089
QY 787 -IATTVAEEGLDIKECNVIRYGLVTNEIAMVQARGBARA-----DESTYVLVAHSGSV 840
Db 1090 EMQAQIAEE--SQIRIELQMTLDSKSDIEQLRSQALHIGLDSSSI-----GSGP 1139
QY 841 IEHETVNDFREKMM--YKAIHCVQNMKPEEYAHKILELQMQSIMCKMKTKRNIKHYKN 898
Db 1140 GDAEADGFPESRLEGWLSLPRVNRNTKKFGWVKYVIVSSKKILFYDSEQDKE-----QS 1194
QY 899 NPSLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKEL---YIVRENKALQKKCADY 954
Db 1195 NPYMVLIDIDLHVRPVTQTDVY-----RADAKEIPRIFQILYANEGESKKEQEF 1244
QY 955 QIN--GE---IICKCGQAWGTMVH----- 974
Db 1245 PVEPVGEKSNYICXKGHEFIPTLYHFPTNCEACMKPLWHMFKPPPALECRRCCHIKCHKDH 1304
QY 975 ----KGLDLPCLJRNFWVVFKN-----NSTKKQYKKWV 1004
Db 1305 MDKKEEIIAPC-KVYYDIDISTAKNLLLLLANSTEEQ-QKWV 1341

```

```

; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 296
; LENGTH: 1388
; TYPE: PR1
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1
US-09-976-594-296

Query Match          3.5%; Score 188.5; DB 4; Length 1388;
Best Local Similarity 18.2%; Pred. No. 8.3e-07;
Matches 204; Conservative 195; Mismatches 399; Indels 321; Gaps 57;

QY 71  WHLGTREFVEALRRGTGSPLAARYMNPETDLPSPSPFENAHDEYQLQLNLLQPTLVDKL 129
Db 359 WH--W-----DNIRETAAPV-----PELSSDIDSSNFDDIEDD---KGDVETFPPIPKA 402

QY 130  LVRDVL-----KCMEEELLTIEDNRIRIAAAENGNESGVRELLKRVQKNWFSAPLNVL 185
Db 403  FVGNQLPFIQFYRENLILSDSP---SCRENDSIQSRKNEESQEIQK-----LYTL 452

QY 186  RQTGNNEL-VQELTGSDCSESNAEINLSQVDGPQVEEQLLSTTVQPNLEKEVWGMENNS 244
Db 453  EEHLSNEMQAKEELEQCKCKSVNTRLEKTAK---ELEEEI---TLRKSVESALRQLEREK 505

QY 245  SESSFADSSVVSSEDTSL-----AEGSVSCLDES�-----GHNSNMGSDS-GTMGSDSD 292
Db 506  ALLQHKNAEYQKADHEADKKNRLENVNSLKDQLEDLKKKNQNSQISTEKVNQLQRQLD 565

QY 293  EENVAARASPEPELQLRPYQMEVAQPA-----LEGKNIICLPTGSGKTRVAVYIA 343
Db 566  ETNALLRTESDAAARLRKTAQESSKQIQOLESNNRDLDKNCL--LETAKLKEKEFINL 623

QY 344  KDHLDDKKKASEPGKIVLVNKKVLLVEQLFRKEFPFLKKYRVIGLSGDTQLKISFPEV 403
Db 624  QSALESERRDRTHGSEII-----NDLQG-----RICGLEED----- 654

QY 404  VKSCDIIISTAQI---LENSLLNLENGEDAGVQLS-DFSLIII-----DECHH----- 447
Db 655  LKNGKILLAKVELEKRLQERFTDLEK-EKSNWEIDMTYQLKVIQQSLEQSEAEHKATKA 713

QY 448  --TNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIPLPQILGLTASPGVGATKQAKAEH 505
Db 714  RLADKNKIYESI-----BEAKSEAMKEMEKKLE-----ERTLKQKVENL 753

QY 506  IL---KLCANLDAFTIKTVKENLDQKNQIQEPCKKFAIADATREDPFKEKLLIEMTRIQ 562
Db 754  LLEAEKRCSLDC-----DLKQSQOKINELLKQ---KDVLNED-VRNLTLEKIEQETQ 801

QY 563  TYCQMSPMDFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHLRKYNEAL-----QI 613
Db 802  KCC--LTQNDLKMQTQQVNTLKMSEKQLQOENNHLMEMKNLEKQNAELRKEKQADGQM 859

QY 614  NDTIRMIDAYTHLETFFYN-----BEKDK-----KFAVIEDSD----- 646
Db 860  KELQDQLEAEQYFSTLYKTQVRELKEECEBTKLKGELQKKQELQDERDSIAAQLIITL 919

QY 647  -----EGGDDEYCDGED-----EDDLKPKL-----KLDETRFLMTLFFENNMKL 687
Db 920  TKADSEQLARSIAEEQYSDLEKEKIMKELEIKEMMARHKQELTEKDATIASL-EETNRTL 978

QY 688  KRLAENPEYENEKLTCLRNTIMEQYTRTEE---SARGIIFTKTROSAYALSQWITEN--- 741
Db 979  TSDVANLANEKEELNNKLKDVOEQLSRLKDDEISAAAAI-----KAOFEKQLLTERTLK 1031

QY 742  ----EKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGTGKINLL----- 786
Db 1032  TQAVNKLAEI-MNRKEPVKFGNDTDVIR-EKEKENRKLHMLKSEREKLTTQOMIKYQKELN 1089

```


; TYPE: PRT		900 PSLITFLCKNSVLACSGEDI--HVIEKMHVNMT-PEFKELYIVRENKALQKKCADYQI 956	
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide			
US-08-973-462-8		1265 KELEKALSSEDSKEIIDAKDDTLEKVIIEEHDITTTLDVVELKDVVEEDKI--EKVSDIKD 1322	
Query Match		957 NGEIICK 963	
Best Local Similarity		1323 LEEDILK 1329	
Matches 201; Conservative 152; Mismatches 373; Indels 301; Gaps 44;			
QY	42	VKEQIORTVATSGNMQAVELLSTLEKGVWHLGWTREFVEALRRITGSPLAARYMNPCLTD 101	
Db	499	VEESVAENVEESVAENVEEIVAPTVEEIV-----APTVEEIVAPSVVE 541	
QY	102	LPSPSFENAHDEYQLNLLQPTLVLDKLVDRDVLDKMEEELLITIEDNRIRIAAENNGN- 160	
Db	542	SVAPSVESVEEN-----VEESVAENVEESVAENVEESVAENVEEII 590	
QY	161	-----ESGVRELLKRIQKENVFSAFLNVLRTGNNELVQELTSGDCSESNAE-----I 209	
Db	591	VAPTVEEIVAPTVEEIVAPSVVESVAPSVVEEEN-----SVAENV 648	
QY	210	ENLSQVDGPQVEEQLLST-----TVQNLKEVWGMENNSSESSFADSSVSES 258	
Db	649	ENVEEIVAPTVEEIVAPTVEEIVAPSVVESVAPSVVEEEN-----SVAENV 699	
QY	259	DTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSEENVAARASPEPELQRPYQMEVAQP 318	
Db	700	EESVAEN-----VEESVAENV-----EESVAENVEEIVAPTVEEIVAP 745	
QY	319	ALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKIVLVNKKVLLVEQLFRKEFQ 378	
Db	746	SV-----VESVAPSVVESVEENV-----EESVAENVEESVAENVEESVA 792	
QY	379	PELKMYRVIGSLDTQLKISFPEVVKSCDIIISTAQILENSLLNLENGEDAGVQLSDFS 438	
Db	793	P-----TVEEIVAPSVVEE-----SVAPSVVEE-----AENVATNLSN 827	
QY	439	LIIIDECHHTNKEAVVNMIMRHYLMQKLKNNRLK-----KENKPVIPLPQILGLTASPGV 493	
Db	828	L-----SNLLGGIETEIEIKDSILNEIEEVKENVVTTILENVEETTAE--- 869	
QY	494	GGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCPKFALADATREDPFKEK 553	
Db	870	-SVTTFNILEEIQE-----NTITNDTIEEKLHELHENVLS-----AALENTQSEERKE 918	
QY	554	LLEIM-----TRIQTVCQMPMS-DFGTOPYEQWAIQMEKKAACKGNRKERVCAEH 603	
Db	919	VIDVIEEVKEEVATTIETVEQAEKKSANTITIFE-----NLEENAVESNEN---VAEN 970	
QY	604	LRKYNEALQINDTIRMIDAYTHL--ETFYNEEKDKKFAVIEDDSDEGGDEYCDG----- 656	
Db	971	LEKINETV-FNTVLDKVEETVEISGESLENNEMDKAFFSEIFDNVKGQENLLTGMFSI 1029	
QY	657	-----DEDEDDLKPLKLDDETRFLMTLFFENNKMRLAENPEYENE-KITKLRNTI 708	
Db	1030	ETSIVIQSEKVDL-----NENVVSSILDNIENMKGLLNKLENI- 1069	
QY	709	MEQYTRTEESARGIIFTKTQSAYALSQWITENEKFAEVGVKA--HLLIG-----AGHSSE 762	
Db	1070	-----SSTEGVQETVTEH-----VEQNVYVDVDPAMKDQFLGILNEAGGLKE 1112	
QY	763	F-----KPMTQNEQKEVTSKFRGTGKINLLIATTVAAEGLDIKE 800	
Db	1113	MFFNLBDFKSESDVITVEEIKDEPVQKEVEKETVSIIEEEN--IVDVLEEKEKDLTD 1170	
QY	801	CNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDFREKMYKAIHC 860	
Db	1171	-----KMIDA-----VEESIEI-----SSDSKEETESIKD-KEKDVSLVVEE 1206	
QY	861	VQNMKPEYAHKILELQMQSIMEKKMK-----TKRNIKHYKNN 899	
Db	1207	VQDNMDDESVEKVLEL--KNMEELMKDAVEINDITSKLIBETQELNEVEADLIKMEKL 1264	
RESULT 13			
US-09-150-867-1			
; Sequence 1, Application US/09150867			
; Patent No. 6645748			
; GENERAL INFORMATION:			
; APPLICANT: Wood, Kenneth W.			
; APPLICANT: Sakowicz, Roman			
; APPLICANT: Goldstein, Lawrence S.B.			
; APPLICANT: Cleveland, Don W.			
; APPLICANT: The Regents of the University of California			
; TITLE OF INVENTION: Plus End-Directed Microtubule Motor Required for			
; TITLE OF INVENTION: Chromosome Congression			
; FILE REFERENCE: 18557C-000110US			
; CURRENT APPLICATION NUMBER: US/09/150,867			
; CURRENT FILING DATE: 1998-09-10			
; EARLIER APPLICATION NUMBER: US 60/058,645			
; EARLIER FILING DATE: 1997-09-11			
; NUMBER OF SEQ ID NOS: 11			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 1			
; LENGTH: 2954			
; TYPE: PRT			
; ORGANISM: Xenopus sp.			
; FEATURE:			
; OTHER INFORMATION: Xenopus centromere-associated protein-E (XCENP-E)			
; OTHER INFORMATION: member of the kinesin superfamily of microtubule			
; OTHER INFORMATION: motor proteins			
; FEATURE:			
; NAME/KEY: DOMAIN			
; LOCATION: (1)..(472)			
; OTHER INFORMATION: kinesin like motor domain			
; FEATURE:			
; NAME/KEY: DOMAIN			
; LOCATION: (473)..(2752)			
; OTHER INFORMATION: rod domain			
; FEATURE:			
; NAME/KEY: DOMAIN			
; LOCATION: (2753)..(2954)			
; OTHER INFORMATION: tail domain			
US-09-150-867-1			
Query Match		3.4%; Score 182.5; DB 4; Length 2954;	
Best Local Similarity		17.9%; Pred. No. 9.2e-06;	
Matches 182; Conservative 187; Mismatches 366; Indels 283; Gaps 42;			
QY	6	STDENFRYLISCFRARKVMIQVEPVLDTFLPAEVEKQIQRTVATSGNMQAVELLST 65	
Db	564	SKEEKKELVQSPFLKI-----AELEEQL-----SVKAKNLEMTVNS 600	
QY	66	LEKGVWHLGWTREFVEALRRITGSPLAARYMNPCLTDLPSPSFE-----NAHDEYLQLNL 120	
Db	601	REHSINAEVQTDVEKEVVRKEMSVLGDGYNASNSDLQDSSVDGKRLSSSHDECIEHRKM 660	
QY	121	LQPTLVD-----KLLVRDVLDKMEEELLITIEDRN---RIAAEN-- 158	
Db	661	LEQKIVDLEEFIEINLKKSENDKQKSSQDFMESIQLCEAIMAEKANALEELMRDNFD 720	
QY	159	----GNESGVREL--LKRIV---QKENWFSAFINVLROTGNNELVQELTSGDCSESNAEI 209	
Db	721	NIILENLTREIADLERSLKENQETNEFEILEKETQKEHEAQLIHEIGSLKLVENAE 780	
QY	210	ENLSQVDGPQVEEQLL-STTVQ-----PNLEKEV-----WGMENNSSESSFADS 252	

Db 781 YNQNLDEEDLETKTKLLKEQEIQLAELRKRADNLQKKVRNFDLSVSMGDSSEKLCEEIFOLK 840
QY 253 SVVSESDTSLAEGSVSC-----LDESIGHNSNMGSDSGTMS-----DSDEENV- 296
Db 841 QSLSDAEAVTRDAQECFLRSNLELKEKMDTNSWYNQKEKAASLEKQLETERSNYK 900
QY 297 AARASPEPELQRLPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDK---KKKA 353
Db 901 KMEADLQKELQSAFNEINYLNGLLAGK-----VPRDILLSRVELEKV 942
QY 354 SEPGKVI-----VLNVKVLVEQLFRKEFQPELKKWYRVIGLSGDTQLKISFPEVVK 405
Db 943 SEFSQLEKALEEKNALENEVTCLSEY---KFLP-----NEVECLKNQISK 985
QY 406 SCDIISTAQIILENSLNLENAGEDAGVQLSDFSLLIIDECHHTNKEAVVNNIMRHYLMQK 465
Db 986 ASEIMLLKQEGEHSASIISKQEIIMQEQSEQILQLTDEVHTQSK-VQQTSEEQYLEMKK 1044
QY 466 LKNNRLKK--ENKPVIPPLQILGLTASPGVGATQAKAEHILKLCANLDAFTIKTVKE 523
Db 1045 MHDDLFEKYIRNK-----SEABD-LLREMENLKG-TMESVEV 1079
QY 524 NLDQLKNQIQEPCCKFAIADATREDPFKEKLE---IMTRIQTQYCOMSPMSDFGTQPYE 579
Db 1080 KIADTKHELEE---TIRD-----KEQLLHEKYYFFQAMQTIPTPLSD----- 1120
QY 580 QWAIQMEKKAACKGNRKERVCAEHLRKYNEALQINDTIRMIDAYTH-----LETFYNE 632
Db 1121 ----SLPPSKLVEGNSQ-----DPIEINDYHNLIALATERNNIMVCLETERNS 1164
QY 633 EKDK-----KFAVIEDDSDSGDDEYCDGDEDEDLKKPLKLDETDRFLMTLFFENNKM 686
Db 1165 LKEQVIDLNTQLSLQAQSIKSDLQPKQDLKEGEVKLLLEME---LLKGHLTDSQLS 1220
QY 687 LKRLAENPEYENEKLTCLRNTIMEQYTR--TEESARGIIFTKTRQSAVALSQWITENEKF 744
Db 1221 IEKL-----QLENLEVTETKLTQLEEMKNITIERNELQTNFEDLKAHDSLKQDLSEN--- 1273
QY 745 AEVGVKAHHLIGAGHSSEFFKPMQNEQKEVSKFR-----PEEYAHKILELQMSIMEKKMK 887
Db 1274 IEQSIETQDELRAAQEE-----LREQQLVDSFRQQLDSCSVGISSPNHDAVANQEKVS 1327
QY 780 TGKINLLIATTVAEEGLDIK-BCNIVIRYGLVTNETIAMVQARGARADESTYVLVAHSGS 838
Db 1328 LGEVNSLQSEMLRGERDELQTSCKALV-----SELELLRAHVKSVEGENLEITKKLNG- 1380
QY 839 GVIEHETVNDFREKMMYKAHCVQNMK-----PEEYAHKILELQMSIMEKKMK 887
Db 1381 --LEKELIGKSESESVLKSM--LENLKEDNNKLKEQAEEYSSKENQFSLEEVPFSGSQK 1434

RESULT 14
US-09-134-001C-5465
; Sequence 5465, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5465
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5465

Query Match 3.4%; Score 179; DB 4; Length 666;
Best Local Similarity 19.5%; Pred. No. 1.7e-06;
Matches 137; Conservative 114; Mismatches 230; Indels 220; Gaps 35;
QY 321 EGKNIIICL-PTGSGKTRVAVYIAKDHLDKKKKASEPGK-VIVLVNKKVLVEQLFR--KE 376
Db 37 EGKRHQTLLGATGTGKTFTMSNVK-----EVGKPTLIIAHNKTLAGQLYSEFKE 86
QY 377 FQPFLLKKWYRVIGLSGDTQLKISFPEV-VKSCDIIISTAQIILENSLNLENAGEDAGVQLS 435
Db 87 FFPENRVEYFV-----SYDYIYQPEAYVPSTDTFIEKASINDEIDQLRHSATSSLFER 140
QY 436 DFSLII--IDECHHTNKEAVVNNIM-----RHYLMQKL-----KNN----- 469
Db 141 DDVILIASVSCIYGLGNPEEYKNLVSVRVGEMERSELLRKLVQVSRNDIDFQRTGF 200
QY 470 RLKKNKPVIPPLQILGLTASPGVGATQAKAEHILKLCANLDAF--TIKTVKENLDQ 527
Db 201 RVRGDDVEIFP-----ASREE---MCIRVEFFGDEIDRIRE-VNY 236
QY 528 LKNQIQEPCCKFAIADA---TREDPFK-----EKLLE----- 556
Db 237 LTGEVIREREHFTTFPASHFVTREKMKVAIERIEKELEERLKELDENKLEAQRLEQR 296
QY 557 -----IMTRIQTQYCOMSPMSDFGTQPY-----EQWAIQMEK-----KA 589
Db 297 TNYDLEMMREMGFCGIENYSVHLTLRPLGSTPYTLDDYFGDDWLVIDESHVTLPOIRG 356
QY 590 AKKGNR-KERVCAEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDEG 648
Db 357 MYNGDRARKQVLIDHGFRLPSALD-NRPLK-----FEF--EETKQLVYV---SATP 403
QY 649 GDDEYCDGDEDEDLKKPLKLDETDRFLMTLFFENNKMMLKRLAENPEYENEKLTCLRNTI 708
Db 404 G-----PYELEHTDEMVEQIIRPTGLDLPKIDVRPT-ENQ-----IDDLL 442
QY 709 MEQVTRTEESARGIIFTKTRQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFFKPMQ 768
Db 443 SEIQDRVDKDERVLVTTLTKMSEDLTTYM-----KEAGIKVNYL-----HSEIKTL-- 489
QY 769 NEQKEVISKFRGTGKINLLIATTVAEEGLDIKECNIVI-----RYGLVTNETIAMVQARG- 822
Db 490 -ERIEIIRDLRMGTYDAIVGINLLREGIDIPESVLVILDADKEGFLRSLSLIQTIGRA 548
QY 823 ARADESTYVL-----VAHSGSGVIEHETVNDFREKMMYKAH 859
Db 549 ARNDKGEVIMYADKITDSMQYAIDETQRRRREIQIAHNKEHGITPKTIN-----KKIH 600
QY 860 CVQNM---KPEEYAHKILELQMSIMEKKMKTKRNIKAHYK 897
Db 601 DVISATVESDETNOQQQTLPKMKTKKERQKTIENTIEKEMK 641

RESULT 15
US-08-685-576-1
; Sequence 1, Application US/08685576
; Patent No. 5906819
; GENERAL INFORMATION:
; APPLICANT: Kaibuchi, Kozo
; APPLICANT: Iwamatsu, Akihiro
; APPLICANT: Nakano, Takeshi
; APPLICANT: Ito, Masaaki
; APPLICANT: Takahashi, No. 5906819uaki
; TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,576
FILING DATE: 24-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-325129
FILING DATE: 05-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-17150
FILING DATE: 05-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-131206
FILING DATE: 26-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16887/843
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-685-576-1

Query Match 3.3%; Score 174; DB 2; Length 1388;
Best Local Similarity 18.5%; Pred. No. 1.5e-05;
Matches 207; Conservative 186; Mismatches 390; Indels 334; Gaps 58;

QY	83	LRRGTGSPLAARYMNP	ELTIEDNRRIAAENNGESVRELLKRIQVKNWFS	AFNLVLRQ	187
Db	364	IRETAAPV	-----PELSSDIDSSNFDDIEDDKGDVETPIPKAFVGNQ	LPFIGTYIRE	418
QY	128	KLLVRDVL	DKMEEELLTI	EDNRRIAAENNGESVRELLKRIQVKNWFS	AFNLVLRQ 187
Db	419	NLLSD	-SPCKEND-SIQSRK-----NEES-QEIQKLYTLEHLS	TEIQ----	461
QY	188	TGNNELVQELT	SGDCSESNAEIENLSQVDGPQVEQLLSTTVQPNLEK	EVGMENSSS	247
Db	462	-AKEELEQK	-----CKSVNTRLEKVAK-----ELEEEI---	TLRKNVESILRLEREKALL	508
QY	248	SFADSSVVS	ESDTSL-----AEGSVSCLDESL-----GHNSNMG	SDS-GTMGSDSDSEN	295
Db	509	QHKNABYQR	KADHEADKKNLNDVNSLKDQLEDLKKRNQNSQISTEK	VNQLQRLDET	568
QY	296	VAARASPEPEL	QRPYQMEVAQPA-----LEGKNIILCLPTGSGKTR	VAVYIAKH	346
Db	569	ALLRTESD	TAARLRKTAESSKQIQQLESNNRDLQDN---CL-----	LETAKLK	615
QY	347	LDKKKKASEP	GKVIVLVNVKVLVVEQLERKEFPFLKKNYRVIGLSG	DTQLKIS-FPEVVK	405
Db	616	LEKE	-----FINLQSVLESERRDRTHGSEII-----	NDLQGRISGLEEDVK	656
QY	406	SCDIIISTAQI	----LENSLLMLENGEDAGVQLS-DFSILII-----	DECH	447
Db	657	NGKILLAKVE	LEKRLQERFTDLEK-EKNMEIDMTYQLKVIQSQSLE	QEETEHTKATKARL	715
QY	448	TNKEAVYNN	IMRHYLMQKLNRLKKNKPKVIPLPQILGLTASPGV	GATKQAKAEHIL	507
Db	716	ADKNKIYESI	-----EAKSEAMKEMEKKLSE-----	ERTLKQKVENLLL	755
QY	508	---	KLCANLDAPTIKTVKENLDQLKNQIQEPCKKFAI	ADATREDPFKEKLEIMTRIQT	564
Db	756	EAEKRC	SILDC-----DLKQSQOKINELLKQ---	KDVLNED-VRNLT	TLKTEQETQKR 803

QY	565	COMSPMSDFGTQPYEQWAIQMEKKA	AKGNRKERVCAEHLRKYNEAL-----QIND	615
Db	804	C--LTQNDLKMOTQOVNTLKNSEK	QOLKQENNHLLEMKMSLEKQNAELRKERQ	DADGQMK 861
QY	616	TIRMIDAYTHLET	FYN-----EEKDK-----KFAVIEDSD	646
Db	862	LQQLAEABQYFSTLYKTQVRELKE	CECKTKLCKELQOKKQELQDERDLSLAAQ	LEITLTK 921
QY	647	-----EGDDDEYCDGDED	-----EDDLKKPL-----KLDETDRFLMTL	FFENNKMVKR 689
Db	922	ADSEQLARSIAE	QYSLEKEKIMKELEIKEMMARHQLTEKD	ATIASL-EETNRLTS 980
QY	690	LAENPEYENEKLT	KLNRNTIMEQVTRTEE---SARGIIFTKTR	QSAYALSQWITEN-----741
Db	981	DVANLANEKEBELN	NNKLKEAQEQLSRLLKDEEISAAAI-----	KAQFEKQLLTERLTKTQ 1033
QY	742	--EKFAEVGVKAHLLIGAGHSSE	FKPMTQNEQKEVISKFRGKINLL-----	I 787
Db	1034	AVNKLAEI-MNRKEPV	KRGNDTDVR-RKEKENRKLHMLKSEREKL	TQQMIKYQKELNEM 1091
QY	788	ATTVAEEGLDIKEC	NIIVIRYGLVTNEIAMVQARGARA-----	DESTYVLVAHSGSGVIE 842
Db	1092	QAQIAEE	-----SQIRIELQWTLDSKSDIEQLRSQ	LQALHGLDSSSI-----GSGPGD 1141
QY	843	HETVNDFREKMM	--YKAIHCVQNMKPEEYAHKILELQMSIME	KMKTKRNIKHYKNP 900
Db	1142	TEADGFPESRLE	GWLSLPVRNNTKKGWVKYVIVSSKKIL	FYDSEQDKE-----QSNP 1196
QY	901	SLITFLCKNCSV	LACSGEDIHVIEKMHVNMTPEFKEL-----	YIVRENKALQKKCADIQI 956
Db	1197	YMVLDIDKLFHVR	PVTQTDVY-----RADAKEIPRIFQIL	LYANEGESKKEQEPV 1246
QY	957	N--GE--IICKGQ	AWGTMVH-----	-----974
Db	1247	EPVGEKSNYICHK	GHEFIPTLYHFPPTNCEACMKPLWHMFK	PPPALCRRCHIKCHDHMD 1306
QY	975	--KGLDLPCLKIR	NFVWFKN-----NSTKKQYKQV	1004
Db	1307	KKEEIIAPC-KVY	YDISSAKNLLLLANSTEEQ-QKWV	1341

Search completed: June 2, 2004, 19:13:05
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:11:53 ; Search time 54 Seconds
(without alignments)
5340.225 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSDDED 1025

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	5311	100.0	1025	14	US-10-055-475-2
2	5311	100.0	1025	14	US-10-228-897-2
3	5307	99.9	1025	14	US-10-055-475-9
4	5299	99.8	1025	14	US-10-055-475-7
5	5285	99.5	1025	16	US-10-275-822A-2
6	2456	46.2	468	15	US-10-094-749-2467
7	1087.5	20.5	925	16	US-10-408-765A-2031
8	1076	20.3	486	14	US-10-055-475-12
9	1040.5	19.6	487	15	US-10-108-260A-4515
10	788.5	14.8	514	14	US-10-055-475-13
11	783	14.7	166	9	US-09-864-761-45372
12	782	14.7	154	14	US-10-106-698-5088
13	782	14.7	154	15	US-10-264-049-2244
14	617	11.6	416	14	US-10-055-475-14
15	563.5	10.6	447	15	US-10-264-049-2303

16	488	9.2	92	9	US-09-864-761-33849	Sequence 33849, A
17	453	8.5	752	15	US-10-369-493-21554	Sequence 21554, A
18	441.5	8.3	650	15	US-10-369-493-1338	Sequence 1338, Ap
19	441.5	8.3	650	15	US-10-369-493-20401	Sequence 20401, A
20	417	7.9	741	15	US-10-369-493-21360	Sequence 21360, A
21	396.5	7.5	485	14	US-10-055-475-17	Sequence 17, Appl
22	384	7.2	549	14	US-10-055-475-16	Sequence 16, Appl
23	357	6.7	71	9	US-09-864-761-47821	Sequence 47821, A
24	342	6.4	738	15	US-10-369-493-1227	Sequence 1227, Ap
25	340	6.4	503	14	US-10-055-475-15	Sequence 15, Appl
26	332	6.3	1924	9	US-09-866-557A-2	Sequence 2, Appli
27	332	6.3	1924	11	US-09-858-862-2	Sequence 2, Appli
28	332	6.3	1924	14	US-10-055-797-2	Sequence 2, Appli
29	332	6.3	1924	16	US-10-350-798-2	Sequence 2, Appli
30	298	5.6	55	9	US-09-864-761-48023	Sequence 48023, A
31	286	5.4	503	13	US-10-027-806-66	Sequence 66, Appl
32	286	5.4	503	13	US-10-034-623-66	Sequence 66, Appl
33	286	5.4	503	14	US-10-027-801-66	Sequence 66, Appl
34	286	5.4	503	14	US-10-029-120-66	Sequence 66, Appl
35	284	5.3	502	13	US-10-027-806-34	Sequence 34, Appl
36	284	5.3	502	13	US-10-034-623-34	Sequence 34, Appl
37	284	5.3	502	14	US-10-027-801-34	Sequence 34, Appl
38	284	5.3	502	14	US-10-029-120-34	Sequence 34, Appl
39	283	5.3	56	9	US-09-864-761-46695	Sequence 46695, A
40	277	5.2	669	15	US-10-104-047-3611	Sequence 3611, Ap
41	244	4.6	50	14	US-10-228-897-10	Sequence 10, Appl
42	214.5	4.0	1790	15	US-10-369-493-1586	Sequence 1586, Ap
43	205.5	3.9	1583	16	US-10-408-765A-1635	Sequence 1635, Ap
44	203	3.8	1905	15	US-10-259-194A-86	Sequence 86, Appl
45	202	3.8	3878	14	US-10-080-608A-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-10-055-475-2
; Sequence 2, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-055-475-2

Query Match 100.0%; Score 5311; DB 14; Length 1025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSGYSTDENFRYLISCFRARKVMIQVEPVDLTFLPAEVKEIQRTVATSGNMQAVE	60
Db	1	MSGYSTDENFRYLISCFRARKVMIQVEPVDLTFLPAEVKEIQRTVATSGNMQAVE	60
QY	61	LLSTLEKGVWHLGTREFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL	120
Db	61	LLSTLEKGVWHLGTREFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL	120
QY	121	LQPTLVKLLVRDVLDKMEEBLITIEDNRNIAAENNGNESGVRELLKRIVQENWFSA	180

```
Db      121 LQPTLVDKLLVRDVLDKMEEELLTIEDNRRIAAAEENNGESGVRELLKRIVQKENWFSA 180
QY      181 FLNVLRTQGNNELVQELTGSDCSSESNAELENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240
Db      181 FLNVLRTQGNNELVQELTGSDCSSESNAELENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240
QY      241 ENNSSESSFADSSVSVSESDTSLAEGSVSCLDSESLGHNSNMGSDSGTMGSDSDDEENVAARA 300
Db      241 ENNSSESSFADSSVSVSESDTSLAEGSVSCLDSESLGHNSNMGSDSGTMGSDSDDEENVAARA 300
QY      301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360
Db      301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360
QY      361 VLNVKVLVVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPPEVVKSCDIIISTAQILENS 420
Db      361 VLNVKVLVVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPPEVVKSCDIIISTAQILENS 420
QY      421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480
Db      421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480
QY      481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
Db      481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
QY      541 IADATREDPFKEKLEIMTRIQTYCQMSPMDFGTQPYEQWAIQMEKKAAGNRKERV 600
Db      541 IADATREDPFKEKLEIMTRIQTYCQMSPMDFGTQPYEQWAIQMEKKAAGNRKERV 600
QY      601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDDEGDDDEYCDGDEDE 660
Db      601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDDEGDDDEYCDGDEDE 660
QY      661 DDLKKPLKLDETDRFLMTLFFENNMKMLKRLAENPEYENEKLTKLRTIMEQYTRTEESAR 720
Db      661 DDLKKPLKLDETDRFLMTLFFENNMKMLKRLAENPEYENEKLTKLRTIMEQYTRTEESAR 720
QY      721 GIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVISKFR 780
Db      721 GIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVISKFR 780
QY      781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840
Db      781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840
QY      841 IEHETVNDFREKMYKAIHCYQNMKPEEYAHKILEIQMSIMEKKMKTKRNIKHYKNP 900
Db      841 IEHETVNDFREKMYKAIHCYQNMKPEEYAHKILEIQMSIMEKKMKTKRNIKHYKNP 900
QY      901 SLITFLCKNCVSLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKKCADYQINGEI 960
Db      901 SLITFLCKNCVSLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKKCADYQINGEI 960
QY      961 ICKGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCL 1020
Db      961 ICKGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCL 1020
QY      1021 FSDED 1025
Db      1021 FSDED 1025
```

RESULT 2
US-10-228-897-2
; Sequence 2, Application US/10228897
; Publication No. US20030092043A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED

```
; TITLE OF INVENTION: GENE-5 AND PROMOTER AND USES THEREOF  
; FILE REFERENCE: A34614-A-PCT-USA (070050.2121)  
; CURRENT APPLICATION NUMBER: US/10/228,897  
; CURRENT FILING DATE: 2002-08-26  
; PRIOR APPLICATION NUMBER: PCT/US01/06960  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 09/515,363  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1025  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-228-897-2
```

```
Query Match      100.0%; Score 5311; DB 14; Length 1025;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY      1 MSGYSTDENFRYLISCERARVKMYIQVEPVLDTLFLPAEVKEQIQRTVATSGNQAVE 60  
Db      1 MSGYSTDENFRYLISCERARVKMYIQVEPVLDTLFLPAEVKEQIQRTVATSGNQAVE 60  
QY      61 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL 120  
Db      61 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL 120  
QY      121 LQPTLVDKLLVRDVLDKMEEELLTIEDNRRIAAAEENNGESGVRELLKRIVQKENWFSA 180  
Db      121 LQPTLVDKLLVRDVLDKMEEELLTIEDNRRIAAAEENNGESGVRELLKRIVQKENWFSA 180  
QY      181 FLNVLRTQGNNELVQELTGSDCSSESNAELENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240  
Db      181 FLNVLRTQGNNELVQELTGSDCSSESNAELENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240  
QY      241 ENNSSESSFADSSVSVSESDTSLAEGSVSCLDSESLGHNSNMGSDSGTMGSDSDDEENVAARA 300  
Db      241 ENNSSESSFADSSVSVSESDTSLAEGSVSCLDSESLGHNSNMGSDSGTMGSDSDDEENVAARA 300  
QY      301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360  
Db      301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360  
QY      361 VLNVKVLVVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPPEVVKSCDIIISTAQILENS 420  
Db      361 VLNVKVLVVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPPEVVKSCDIIISTAQILENS 420  
QY      421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480  
Db      421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480  
QY      481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540  
Db      481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540  
QY      541 IADATREDPFKEKLEIMTRIQTYCQMSPMDFGTQPYEQWAIQMEKKAAGNRKERV 600  
Db      541 IADATREDPFKEKLEIMTRIQTYCQMSPMDFGTQPYEQWAIQMEKKAAGNRKERV 600  
QY      601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDDEGDDDEYCDGDEDE 660  
Db      601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDDEGDDDEYCDGDEDE 660  
QY      661 DDLKKPLKLDETDRFLMTLFFENNMKMLKRLAENPEYENEKLTKLRTIMEQYTRTEESAR 720  
Db      661 DDLKKPLKLDETDRFLMTLFFENNMKMLKRLAENPEYENEKLTKLRTIMEQYTRTEESAR 720  
QY      721 GIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVISKFR 780  
Db      721 GIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVISKFR 780  
QY      781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840
```

Db 781 GKNLLIATVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRADESTYVLVAHSGSGV 840
QY 841 IEHETVNDFREKMMYKAHCVQNMKPEEYAHKILELOMQSIMCKMKTKRNIKHYKNP 900
Db 841 IEHETVNDFREKMMYKAHCVQNMKPEEYAHKILELOMQSIMCKMKTKRNIKHYKNP 900
QY 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALOKKCADYQINGEI 960
Db 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALOKKCADYQINGEI 960
QY 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCL 1020
Db 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCL 1020
QY 1021 FSDED 1025
Db 1021 FSDED 1025
RESULT 3
US-10-055-475-9
; Sequence 9, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-055-475-9
Query Match 99.9%; Score 5307; DB 14; Length 1025;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1024; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSGYSTDENFRYLISCFRARKVMIQVEPVDLYTLFLPAEVEKEIQRTVATSGNMQAVE 60
Db 1 MSGYSTDENFRYLISCFRARKVMIQVEPVDLYTLFLPAEVEKEIQRTVATSGNMQAVE 60
QY 61 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL 120
Db 61 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL 120
QY 121 LQPTLVDKLLVRDVLDKMEEELLTIEDNRNIAAENNGNMGVRELLKRIVQKENWFSA 180
Db 121 LQPTLVDKLLVRDVLDKMEEELLTIEDNRNIAAENNGNMGVRELLKRIVQKENWFSA 180
QY 181 FLNVLROTGNNELVQELTGSDCSESNAEIEENLSQVDGPQVEEQLSTTVQPNLEKEVWGM 240
Db 181 FLNVLROTGNNELVQELTGSDCSESNAEIEENLSQVDGPQVEEQLSTTVQPNLEKEVWGM 240
QY 241 ENNSSESSFADSSVVSSEDTSLAEGSVCSLDESIGHNSNMGSDSGTMGSDSDEENVAARA 300
Db 241 ENNSSESSFADSSVVSSEDTSLAEGSVCSLDESIGHNSNMGSDSGTMGSDSDEENVAARA 300
QY 301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
Db 301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360

QY 361 VLVNKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISFPEVVVKSCDIIISTAILENS 420
Db 361 VLVNKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISFPEVVVKSCDIIISTAILENS 420
QY 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPKVIP 480
Db 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPKVIP 480
QY 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIOEPCCKFA 540
Db 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIOEPCCKFA 540
QY 541 IADATREDPEKEKLLIEMTRIQTQYCOMSPMSDFGTOPYEQWAIQMEKKAACKGNRKERV 600
Db 541 IADATREDPEKEKLLIEMTRIQTQYCOMSPMSDFGTOPYEQWAIQMEKKAACKGNRKERV 600
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKKKFAVIEDDSDEGGDDDEYCDGDEDE 660
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKKKFAVIEDDSDEGGDDDEYCDGDEDE 660
QY 661 DDLKKPLKLDDETRFLMTLFFENNMKMLKRLAENPEYENEKTLKLRNTIMEQYTRTEESAR 720
Db 661 DDLKKPLKLDDETRFLMTLFFENNMKMLKRLAENPEYENEKTLKLRNTIMEQYTRTEESAR 720
QY 721 GIIFTKTRQSAYALSQWITENKFAEVGVKAHHLIGAGHSSEFPKMTQNEQKEVISKFT 780
Db 721 GIIFTKTRQSAYALSQWITENKFAEVGVKAHHLIGAGHSSEFPKMTQNEQKEVISKFT 780
QY 781 GKNLLIATVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRADESTYVLVAHSGSGV 840
Db 781 GKNLLIATVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRADESTYVLVAHSGSGV 840
QY 841 IEHETVNDFREKMMYKAHCVQNMKPEEYAHKILELOMQSIMCKMKTKRNIKHYKNP 900
Db 841 IEHETVNDFREKMMYKAHCVQNMKPEEYAHKILELOMQSIMCKMKTKRNIKHYKNP 900
QY 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALOKKCADYQINGEI 960
Db 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALOKKCADYQINGEI 960
QY 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCL 1020
Db 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCL 1020
QY 1021 FSDED 1025
Db 1021 FSDED 1025
RESULT 4
US-10-055-475-7
; Sequence 7, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-055-475-7

Query Match		99.8%;	Score 5299;	DB 14;	Length 1025;
Best Local Similarity		99.8%;	Pred. No. 0;		
Matches 1023;		Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MSNGYSTDENFRYLISCFRARKVMIYQVEPVLDYLTFLPAEVKEQIQRTVATSGNQAVE	60		
Db	1	MSNGYSTDENFRYLISCFRARKVMIYQVEPVLDYLTFLPAEVKEQIQRTVATSGNQAVE	60		
Qy	61	LLSLTLEKGVVHGLWTREFVEALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL	120		
Db	61	LLSLTLEKGVVHGLWTREFVEALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL	120		
Qy	121	LQPTLVDKLLVRDVLDKCMEEBELLTIEDNRRIAAENNGNESGVRELLKRIVQKENWFS	180		
Db	121	LQPTLVDKLLVRDVLDKCMEEBELLTIEDNRRIAAENNGNESGVRELLKRIVQKENWFS	180		
Qy	181	FLNVLRQTGNNELVQELTGDCSSENAEIEIENLSQVDPQVVEQLLSTTVQPNLEKEVWGM	240		
Db	181	FLNVLRQTGNNELVQELTGDCSSENAEIEIENLSQVDPQVVEQLLSTTVQPNLEKEVWGM	240		
Qy	241	ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSEENVAARA	300		
Db	241	ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSEENVAARA	300		
Qy	301	SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDDKKKASEPGKVI	360		
Db	301	SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDDKKKASEPGKVI	360		
Qy	361	VLNVKVLVLEQLFRKEFQPLKKWYRVIGLSGDTQLKISFPFVVKSCDIIISTAQILENS	420		
Db	361	VLNVKVLVLEQLFRKEFQPLKKWYRVIGLSGDTQLKISFPFVVKSCDIIISTAQILENS	420		
Qy	421	LLNLENGEDAGVQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKENKPVIP	480		
Db	421	LLNLENGEDAGVQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKENKPVIP	480		
Qy	481	LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA	540		
Db	481	LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA	540		
Qy	541	IADATREDPFKEKLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGNRRKERV	600		
Db	541	IADATREDPFKEKLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGNRRKERV	600		
Qy	601	AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDSDEGDDDEYCDGEDE	660		
Db	601	AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDSDEGDDDEYCDGEDE	660		
Qy	661	DDLKKPLKLDLTDTRFMTLFFENNMKMLKRLAENPEYENKLTCLRNTIMEQYTRTEESAR	720		
Db	661	DDLKKPLKLDLTDTRFMTLFFENNMKMLKRLAENPEYENKLTCLRNTIMEQYTRTEESAR	720		
Qy	721	GIIFTKTQRSAYALSQWITENEKFAEYGVVKAHHLIGAGHSSEFKPMTQNEQKEVISKFR	780		
Db	721	GIIFTKTQRSAYALSQWITENEKFAEYGVVKAHHLIGAGHSSEFKPMTQNEQKEVISKFR	780		
Qy	781	GKINLLIATTVAEEGLDIKECNIVIRYGLVTNETIAMVQARGRARESTYVLVAHSGSV	840		
Db	781	GKINLLIATTVAEEGLDIKECNIVIRYGLVTNETIAMVQARGRARESTYVLVAHSGSV	840		
Qy	841	IEHETVNDFREKMYKAIHCYQNMKPEEYAHKILBELQMQSIMKKMTKRNIAKHYKNP	900		
Db	841	IERETVNDFREKMYKAIHCYQNMKPEEYAHKILBELQMQSIMKKMTKRNIAKHYKNP	900		
Qy	901	SLITFLCKNCSVLACSGEDIHVEIKMHVNMTPPEFKELYIVRENKALQKCADYQINGEI	960		
Db	901	SLITFLCKNCSVLACSGEDIHVEIKMHVNMTPPEFKELYIVRENKALQKCADYQINGEI	960		
Qy	961	ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCL	1020		
Db	961	ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCL	1020		

QY	1021	FSDED	1025		
Db	1021	FSDED	1025		
RESULT 5					
US-10-275-822A-2					
; Sequence 2, Application US/10275822A					
; Publication No. US20040086500A1					
; GENERAL INFORMATION:					
; APPLICANT: Bahr, Georges					
; APPLICANT: Cocude, Cecile					
; APPLICANT: Capron, Andre					
; TITLE OF INVENTION: RH16 Polypeptide and its Fragments and Polynucleotides Encoding					
; TITLE OF INVENTION: said Polypeptides and Therapeutic Uses					
; FILE REFERENCE: 017753-170					
; CURRENT APPLICATION NUMBER: US/10/275,822A					
; CURRENT FILING DATE: 2002-11-12					
; PRIOR APPLICATION NUMBER: PCT/FR01/01441					
; PRIOR FILING DATE: 2001-05-11					
; PRIOR APPLICATION NUMBER: FR 00/06,030					
; PRIOR FILING DATE: 2000-05-11					
; NUMBER OF SEQ ID NOS: 56					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 2					
; LENGTH: 1025					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-10-275-822A-2					
Query Match 99.5%; Score 5285; DB 16; Length 1025;					
Best Local Similarity 99.5%; Pred. No. 0;					
Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;					
QY	1	MSNGYSTDENFRYLISCFRARKVMIYQVEPVLDYLTFLPAEVKEQIQRTVATSGNQAVE	60		
Db	1	MSNGYSTDENFRYLISCFRARKVMIYQVEPVLDYLTFLPAEVKEQIQRTVATSGNQAVE	60		
QY	61	LLSLTLEKGVVHGLWTREFVEALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL	120		
Db	61	LLSLTLEKGVVHGLWTREFVEALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL	120		
QY	121	LQPTLVDKLLVRDVLDKCMEEELLTIEDNRRIAAENNGNESGVRELLKRIVQKENWFS	180		
Db	121	LQPTLVDKLLVRDVLDKCMEEELLTIEDNRRIAAENNGNESGVRELLKRIVQKENWFS	180		
QY	181	FLNVLRQTGNNELVQELTGDCSSENAEIEIENLSQVDPQVVEQLLSTTVQPNLEKEVWGM	240		
Db	181	FLNVLRQTGNNELVQELTGDCSSENAEIEIENLSQVDPQVVEQLLSTTVQPNLEKEVWGM	240		
QY	241	ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSEENVAARA	300		
Db	241	ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSEENVAARA	300		
QY	301	SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDDKKKASEPGKVI	360		
Db	301	SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDDKKKASEPGKVI	360		
QY	361	VLNVKVLVLEQLFRKEFQPLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS	420		
Db	361	VLNVKVLVLEQLFRKEFQPLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS	420		
QY	421	LLNLENGEDAGVQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKENKPVIP	480		
Db	421	LLNLENGEDAGVQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKENKPVIP	480		
QY	481	LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA	540		
Db	481	LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA	540		
QY	541	IADATREDPFKEKLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGNRRKERV	600		
Db	541	IADATREDPFKEKLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGNRRKERV	600		

QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKOKKFAVIEDDSDEGGDEYCDGDEDE 660
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKOKKFAVIEDDSDEGGDEYCDGDEDE 660
QY 661 DDLKKPLKLDDETRFLMTLFFENNMKRLAENPEYENEXLTKLRNTIMEQYTRTEESAR 720
Db 661 DDLKKPLKLDDETRFLMTLFFENNMKRLAENPEYENEXLTKLRNTIMEQYTRTEESAR 720
QY 721 GIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRT 780
Db 721 GIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRT 780
QY 781 GKINLLIATTVABEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840
Db 781 GKINLLIATTVABEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840
QY 841 IEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMSIMEKKMKTKRNIACHYKNNP 900
Db 841 IERETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMSIMEKKMKTKRNIACHYKNNP 900
QY 901 SLITFLCKNCSVLACSGEDIHVIERKMHVNMTPFEKELYIVRENKALQKKCADYQINGEI 960
Db 901 SLITFLCKNCSVLACSGEDIHVIERKMHVNMTPFEKELYIVRENKALQKKCADYQINGEI 960
QY 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVFVFNKSTKKQYKKWVELPITFPNLDYSECCL 1020
Db 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVFVFNKSTKKQYKKWVELPITFPNLDYSECCL 1020
QY 1021 FSDED 1025
Db 1021 FSDED 1025

RESULT 6
US-10-094-749-2467
; Sequence 2467, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NACHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2467
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2467

Query Match 46.2%; Score 2456; DB 15; Length 468;
Best Local Similarity 99.1%; Pred. No. 9e-171;
Matches 464; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 558 MTRIQTQYQMSPMDFGTQPYEQWAIQMEKKAAGKGNKRVCAEHLRKYNEALQINDTI 617
Db 1 MTRIQTQYQMSPMDFGTQPYEQWAIQMEKKAAGKGNKRVCAEHLRKYNEALQINDTI 60
QY 618 RMIDAYTHLETFYNEEKOKKFAVIEDDSDEGGDEYCDGDEDEDDLLKKPLKLDDETRFLM 677
Db 61 RMIDAYTHLETFYNEEKOKKFAVIEDDSDEGGDEYCDGDEDEDDLLKKPLKLDDETRFLM 120
QY 678 TLFFENNMKRLAENPEYENEXLTKLRNTIMEQYTRTEESARGIIFTKTRQSAYALSQW 737
Db 121 TLFFENNMKRLAENPEYENEXLTKLRNTIMEQYTRTEESARGIIFTKTRQSAYALSQW 180
QY 738 ITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRTGKINLLIATTVABEGLD 797
Db 181 ITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRTGRINLLIATTVABEGLD 240
QY 798 IKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDFREKMYKA 857
Db 241 IKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIERETVNDFREKMYKA 300
QY 858 IHCVNMPPEYAHKILELQMSIMEKKMKTKRNIACHYKNNPSLITFLCKNCSVLACSG 917
Db 301 IHCVNMPPEYAHKILELQMSIMEKKMKTKRNIACHYKNNPSLITFLCKNCSVLACSG 360
QY 918 EDIHVIERKMHVNMTPFEKELYIVRENKALQKKCADYQINGEIIICKGQAWGTMVHKGL 977
Db 361 EDIHVIERKMHVNMTPFEKELYIVRENKALQKKCADYQINGEIIICKGQAWGTMVHKGL 420
QY 978 DLPCLKIRNFVVFVFNKSTKKQYKKWVELPITFPNLDYSECCLFSDED 1025
Db 421 DLPCLKIRNFVVFVFNKSTKKQYKKWVELPITFPNLDYSECCLFSDED 468

RESULT 7
US-10-408-765A-2031
; Sequence 2031, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2031
; LENGTH: 925
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2031

Query Match 20.5%; Score 1087.5; DB 16; Length 925;
Best Local Similarity 31.5%; Pred. No. 2.5e-70;
Matches 328; Conservative 166; Mismatches 361; Indels 185; Gaps 37;
QY 12 RYLISCFRVRKMYIQVEPVDYLT--FLPAEVKEQIQRTVATSGNMQAVELLSTLEKG 69
Db 6 RRSQAQDYIRKTLDPITYLSYMAPWFRREEV-QYIAEKNNKGPMEATLFLKFLLE- 63
QY 70 VVHLGWTFEFVEALRRRTGSLAARYMNPDLTDLSPSPFENAH--DEYQLNLLQPTLVD 127
Db 64 LQEEGWFRGFLDALDHAG--YSGLYEAIESWD-----FKKIEKLEEYRLLKRLQPEFKT 116


```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4515

Query Match      19.6%; Score 1040.5; DB 15; Length 487;
Best Local Similarity 41.1%; Pred. No. 2.6e-67;
Matches 234; Conservative 90; Mismatches 161; Indels 85; Gaps 9;

QY 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKXKASEPGKVIVLVNK 365
Db 1 MELRSYQWEVIMPALEGKNIICLPTGAGKTRAAAYVAKRHLE---TVDGAKVVVLVNR 56

QY 366 VLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENSLNL 425
Db 57 VHLVTQ-HGEFRRMLDGRWTVTTLSGDMGPRAGFGHLAR----- 95

QY 426 NGEDAGVQLSDFSIIIDECHHTNKEAVYNNIMRHVLMQKLNKLNKKNKPKVIPLPQIL 485
Db 96 -----CHHTKDTVNVNIMSQYLELKLQRAQ-----PLPQVL 127

QY 486 GLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQKNQIQEPCPKFAIADAT 545
Db 128 GLTASPGTGASKLDGAINHVQLCANLDTWCIMSPQNCPCQLQEHSSQPCQYNLCHRR 187

QY 546 REDPFKEKLLLEIMTRIQTQYCOMSPMS-DEGTQPYEQWAIQMEKKAAGKGNRKERVCAEHL 604
Db 188 SQDPFGDLLKXLMQDIHDHLEMPELSRKFGTQMYEQVVVKLSEAAALAGLQEQRVYALHL 247

QY 605 RYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDDLK 664
Db 248 RRYNDALLIHDIVRAVDALAAALQDFYHREHVTKTQIL-----C----- 285

QY 665 KPLKLDETDRFLMTLFFENNMKMLKRLAENPEYENEKLTKLRTIMEQYTRTEESARGIIF 724
Db 286 -----AERRLLALFDDRKNELAHATHGP-ENPKLEMLEKILQRFSS-SSNSPRGIIF 336

QY 725 TKTRQSAYALSQWITENEKFAEVGVVKAHHLIGAGHSSEFKPMQNEQKEVISKFRICKIN 784
Db 337 TRTRQSAHSLLLWLOQQOGLQTVDIRAQLLIGAGNSSQSTMTQDQEQEVIQKFDGTLN 396

QY 785 LLIATTVAAEGLDIKECNIVIRYGLVTVNEIAMVQARGRARADESTYVLVAHSGSGVIEHE 844
Db 397 LLVATSVAAEGLDIPHNCVVVRYGLLTNEISMVQARGRARAQSQSVYAFVATEGSRLEKRE 456

QY 845 TVNDFREKMYKAHCVQNMKPEEYAKIL 874
Db 457 LINEALETLMQAVAAVQKMDQAEYQAKEL 486
```

```

RESULT 10
US-10-055-475-13
; Sequence 13, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 514
; TYPE: PRT
; ORGANISM: sus scrofa
US-10-055-475-13
```

```

Query Match      14.8%; Score 788.5; DB 14; Length 514;
Best Local Similarity 37.4%; Pred. No. 7.3e-49;
Matches 210; Conservative 89; Mismatches 178; Indels 85; Gaps 16;

QY 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKXKASEPGKVIVLVNK 365
Db 6 LKPRKYQLELALPAQNGKNTIICAPTGCCTFVSLICEHHL-KKPPRGRKGKVVFFAIQ 64

QY 366 VLLVEQ---LFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENSLL 422
Db 65 LPVYEQQKSVFSKHFE---RLGYKVAGISGATSDTVCVEQIVENSDDIILTPQILVNCLT 121

QY 423 NLENGEDAGV--QLSDFSIIIDECHHTNKEAVYNNIMRHVLMQKLNKLNKKNKPKVIP 480
Db 122 N-----GTIPSLSVFTLMIFDECHNTSKQHPYNVIMFSYLDKLGSS-----DS 166

QY 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQKNQIQEPCPKFA 540
Db 167 LPQVIGLTASVGDGAKNKAETEYICKLCASLDTSVIATVRDNLLEELEEVVYKPKFFR 226

QY 541 IADATREDPFKEKLLLEIMTRIQTQYCOMSPMS-DEGTQPYEQWAIQMEKKAAGKGNRK 585
Db 227 KVELRTTDRFKCIISQLMMEIESLAKSIPEELGTITLGLFQIQNSNFGTQKYEQWIVKV 286

QY 586 EKAA-----KGNRKERVCA-DEGTQPYEQWAIQMEKKAAGKGNRKERVCAEHL 635
Db 287 QKECAVFPMPDKDKESRICKALFSYMSHLRIYNDALIINEHARMKDALDYLKOFFRNIRA 346

QY 636 KKFVAVIEDDSDEGGDDEYCDGDEDDLKPLKLDETDRFLMTLFFENNMKMLKRLAENPE 695
Db 347 AGF-----DEIEQDL-----TQR-----FEELQLELESISIDPS 375

QY 696 YENEKLTKLRTIMEQYTRTEESARGIIFTKTRQSAYALSQWITENEKFAEVGVVKAHHLI 755
Db 376 NENPKLRDLCLFILQEEYHLNPEF-RTILFVKTRALVDALKKIKENPKLS--FLKPSILT 432

QY 756 GAGHSSEFKPMQNEQKEVISKFRICKIN-LLIATTVAAEGLDIKECNIVIRYGLVTVNEI 814
Db 433 GRGKTNQNIQMTLPQKCVLDTFRDKDKILITTSVADEGIDIAQCNLVILYEVGVNVI 492

QY 815 AMVQARGRARADES-TYVLVAH 835
Db 493 KMIQTRGRGRARGSKCFLLTAN 514
```

```

RESULT 11
US-09-864-761-45372
; Sequence 45372, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45372
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010876.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.58
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.5
; OTHER INFORMATION: EST HUMAN HIT: A1906351.1, EVALUE 1.00e-42
; OTHER INFORMATION: SWISSPROT HIT: P14294, EVALUE 2.20e+00
US-09-864-761-45372

Query Match 14.7%; Score 783; DB 9; Length 166;
Best Local Similarity 93.3%; Pred. No. 3.3e-49;
Matches 152; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSNGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTFLPAEVKEQIQRTVATSGNMQAVE 60
Db 1 MSNGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTFLPAEVKEQIQRTVATSGNMQAVE 60

QY 61 LLLSTLEKGVHGLWGTREFVEALRRTCGSPLAARYMNPDLTLPSPSFENAHDEYQLLNL 120
Db 61 LLLSTLEKGVHGLWGTREFVEALRRTCGSPLAARYMNPDLTLPSPSFENAHDEYQLLNL 120

QY 121 LQPTLVDKLLVRDVLDKCMEEEELLTTIEDNRNRIAAENNGESG 163
Db 121 LQPTLVDKLLVRDVLDKCMEEEELLTTIEDNRNRVGVCSGASLLG 163

RESULT 12
US-10-106-698-5088
; Sequence 5088, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564

; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5088
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (11)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (23)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-10-106-698-5088

Query Match 14.7%; Score 782; DB 14; Length 154;
Best Local Similarity 94.8%; Pred. No. 3.5e-49;
Matches 145; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 873 ILELQMSIMEKQMTKRNIAKHYKNPNSLITFLCKNCVLCACSGEDIHVIEKMHVNM 932
Db 2 ILELQMSIXEKMKXXRNIAKHXNXPNSLITFLCKNCVLCACSGEDIHVIEKMHVNM 61

QY 933 PEFKELYIVRENKALQKCADYQINGEIIICKGQAWGTMVHKGLDLPCLKIRNFVVVK 992
Db 62 PEFKELYIVRENKXLOKCADYQINGEIIICKGQAWGTMVHKGLDLPCLKIRNFVVVK 121

QY 993 NNSTKKQYKKWVELPITFPNLDYSECCLFSD 1025
Db 122 NNSTKKQYKKWVELPITFPNLDYSECCLFSD 154

RESULT 13
US-10-264-049-2244
; Sequence 2244, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2244
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (11)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

FEATURE: MISC_FEATURE
NAME/KEY: MISC_FEATURE (17)
LOCATION: (17)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE: MISC_FEATURE
NAME/KEY: MISC_FEATURE (18)
LOCATION: (18)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE: MISC_FEATURE
NAME/KEY: MISC_FEATURE (23)
LOCATION: (23)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE: MISC_FEATURE
NAME/KEY: MISC_FEATURE (25)
LOCATION: (25)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE: MISC_FEATURE
NAME/KEY: MISC_FEATURE (26)
LOCATION: (26)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE: MISC_FEATURE
NAME/KEY: MISC_FEATURE (28)
LOCATION: (28)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE: MISC_FEATURE
NAME/KEY: MISC_FEATURE (75)
LOCATION: (75)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-2244

Query Match 14.7%; Score 782; DB 15; Length 154;
Best Local Similarity 94.8%; Pred. No. 3.5e-49;
Matches 145; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 873 ILELQMSIMEKMKTKRNIAXHKKXNPSLITFLCKNCSVLACSGEDIHVIKMHVNMT 932
Db 2 ILELQMSIXEKMKXXRNIAHXXNPSLITFLCKNCSVLACSGEDIHVIKMHVNMT 61
QY 933 PEFKELYIVRENKALQKKCADYQINGELICKGQAWGTMVHKGLDLPCLKIRNFVVVFK 992
Db 62 PEFKELYIVRENKXLLQKKCADYQINGELICKGQAWGTMVHKGLDLPCLKIRNFVVVFK 121
QY 993 NNSTKKQYKXWVLPITFPNLDYSECLFSD 1025
Db 122 NNSTKKQYKXWVLPITFPNLDYSECLFSD 154

RESULT 14
US-10-055-475-14
Sequence 14, Application US/10055475
Publication No. US20030022855A1
GENERAL INFORMATION:
APPLICANT: Fisher, Paul B.
APPLICANT: Kang, Dong-Chul
APPLICANT: Gopalkrishnan, Rahul V.
TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
CURRENT APPLICATION NUMBER: US/10/055.475
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: PCT/US01/06960
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 09/515,363
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 416
TYPE: PRT
ORGANISM: homo sapiens
US-10-055-475-14

Query Match 11.6%; Score 617; DB 14; Length 416;

Best Local Similarity 32.8%; Pred. No. 1.8e-36;
Matches 174; Conservative 77; Mismatches 148; Indels 132; Gaps 15;
QY 309 RPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVIVLVNKVLL 368
Db 9 RNYQLELALPAWKGNITICAPTGCFTFVSLICEHHL-KKPPQGGKGVFFANQIPV 67
QY 369 VEQ---LFRKEFPFLKKWYRVIGLSDTQLKISFPEVVKSCDIIISTAQILENSLLNLE 425
Db 68 YEQQKSVFSKYFE---RHGYRVTGISGATAENVPVEQIVENNIIILTPQILVN---NLK 121
QY 426 NGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNKNNRLKKNKPVIPLPQIL 485
Db 122 KGTP--SLSIFTLMI FDECHNTSKQHPYNNMIMFNLYDQKLGSS-----GPLPQVI 171
QY 486 GLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDOLKNQIQEPCCKFAIADAT 545
Db 172 GLTASVGVGDAKNTDEALDYICKLCAV---IATVKNLELEQVYVKPKQFFRKVESR 227
QY 546 REDPFEKELLEIMTRIQTCCQSPMSDFGTQPYEQWAIQMEKKAAGKGNRKERVCAHLR 605
Db 228 ISDKFYIIAQLMRDTESLA-----KRIKCD--- 253
QY 606 KYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSDEGGDDEYCDGEDDLKK 665
Db 254 -----LENLSQIONRE----- 264
QY 666 PLKLDTRFLMTLFFENNMKRLAENPEYENKLTKLNTIMEQYTRTEESARGIIFT 725
Db 265 --KLQE-----LESVSRDPSNENPKLEDLCFILQEEYHLNPETIT-ILFV 306
QY 726 KTRQSAVALSQWITENEK--FAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVISKFR-TGK 782
Db 307 KTRALVDALKNWIEGNPKLSFLKPGI---LTGRGKTQNTGMTPLPAQKCILDFAKASGD 362
QY 783 INLLIATVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLV 833
Db 363 HNILIATSVADGIDIAQCNLVILYEVVGNVIXMIQTRGRGRGSKCFL 413

RESULT 15
US-10-264-049-2303
Sequence 2303, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA133P1
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2303
LENGTH: 447
TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-049-2303

Query Match 10.6%; Score 563.5; DB 15; Length 447;
Best Local Similarity 33.0%; Pred. No. 1.6e-32;
Matches 155; Conservative 75; Mismatches 170; Indels 69; Gaps 16;
QY 553 KLEIMTRIQTCCQSPMSDFGTQPYEQWAIQMEKKA---AKGNRKERV---AE 602
Db 13 KLENLSQION-----REFGTQKYEQWIVTVQKACMVQFQMPDKDESRICKALFLYTS 65
QY 603 HLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSDEGGDDEYCDGEDDD 662
Db 66 HLRKYNDALIISEHARMKDALDYKDFFSNVRAAGF-----DEIEQD 107

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:07:33 ; Search time 27 Seconds
(without alignments)
3651.715 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGGYSTDENFRYLISCRA.....LPITFPNLDYSECCLFSDDED 1025

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	687	12.9	1037	2 T32534	hypothetical prote
2	654	12.3	811	2 T30968	helicase homolog C
3	642.5	12.1	956	2 A88708	protein C01B10.1 [
4	516	9.7	398	2 T46312	hypothetical prote
5	499	9.4	1134	2 T20332	hypothetical prote
6	453	8.5	752	2 D75219	ATP-dependent RNA
7	441.5	8.3	650	2 D71203	probable ATP-depen
8	417	7.9	741	1 A69432	ATP-dependent RNA
9	402.5	7.6	1374	2 S62524	probable RNA helic
10	392.5	7.4	784	2 H64487	eIF-4A family prob
11	350	6.6	1822	2 S44849	K12H4.8 protein -
12	342	6.4	738	1 E69055	ATP-dependent RNA
13	313.5	5.9	784	2 D84386	ATP-dependent RNA
14	303.5	5.7	993	2 S48436	probable RNA helic
15	284	5.3	502	2 T31323	ATP-dependent RNA
16	278.5	5.2	839	2 T39190	probable ATP-depen
17	223	4.2	557	2 G69958	SNF2 helicase homo
18	218	4.1	1979	2 C71622	hypothetical prote
19	217.5	4.1	2663	1 S28261	centromere protein
20	214.5	4.0	1790	2 S67593	transport protein
21	209.5	3.9	449	2 C71156	probable helicase
22	205	3.9	453	2 B75006	DNA repair protein
23	205	3.9	967	2 D72308	conserved hypothet
24	204.5	3.9	2163	2 S50675	pre-mRNA splicing
25	204	3.8	1939	2 T18372	repeat organellar
26	201.5	3.8	1063	2 T38420	probable DNA helic
27	198.5	3.7	2253	2 T30336	nuclear/mitotic ap
28	197.5	3.7	1163	2 G97236	ATPase involved in
29	195	3.7	1531	2 T48946	hypothetical prote

30	195	3.7	3259	1 A56539	giantin - human
31	194.5	3.7	986	2 E90596	restriction-modifi
32	194.5	3.7	1256	2 T26101	hypothetical prote
33	194	3.7	1039	2 E72734	hypothetical prote
34	193.5	3.6	663	2 G89848	exinuclease ABC su
35	193.5	3.6	2017	1 A36014	myosin heavy chain
36	193.5	3.6	2057	2 S61477	myosin II heavy ch
37	193.5	3.6	3225	2 I52300	giantin - human
38	193	3.6	715	2 B75135	DNA helicase relat
39	193	3.6	986	2 H90565	restriction modifi
40	192.5	3.6	718	2 E71100	hypothetical prote
41	192.5	3.6	1300	2 I53799	CG1 protein - huma
42	192	3.6	986	2 S49394	HedR1 protein - My
43	192	3.6	1356	2 S32763	kinectin 1 - human
44	191.5	3.6	943	2 F69543	ATP-dependent RNA
45	191	3.6	1922	2 T00637	hypothetical prote

ALIGNMENTS

RESULT 1

T32534

hypothetical protein F15B10.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C;Accession: T32534

R;Wohldmann, P.; Murray, J.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of C. elegans cosmid F15B10.

A;Reference number: Z21186

A;Accession: T32534

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1037 <WOH>

A;Cross-references: EMBL:AF036696; PIDN:AAB88350.1; GSPDB:GN00022; CESP:F15B10.2

A;Experimental source: strain Bristol N2; clone F15B10

C;Genetics:

A;Gene: CESP:F15B10.2

A;Map position: 4

A;Introns: 30/3; 89/1; 111/2; 218/1; 258/2; 288/1; 343/3; 370/3; 400/2; 443/1; 520/1; 55

Query Match		12.9%;	Score 687;	DB 2;	Length 1037;
Best Local Similarity		25.0%;	Pred. No. 1.4e-27;		
Matches 251;		Conservative 169;	Mismatches 408;	Indels 174;	Gaps 35;
QY	101	DLPSPSFENAHDE--YLQLLNLLQPTLVLDKLLVRDVLDKMEEE-----	142		
Db	63	DLKNSILSNADDERLYKDIMTVLQ-TYLPKCTVHKLLN-CSNREVKLSDFHYLDHFEFG	120		
QY	143	LLTIEDNRNRIAAENNGNESGVRELLKRIVQK--ENWFSAPLN--VLQRTG--NNELVQ	195		
Db	121	LRFIEPKVVLAYLDSYPQYIDAVAVLRKEIERNEEDNQDSDFIKKILRTVPLLGEQAVY	180		
QY	196	EL--TGSDCSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGMEN-----N	243		
Db	181	DIMYTISEKSSNNLDVE-----AKQFIA-----KVLRLKNDGFLRFYQIIN	221		
QY	244	SSESSFADSSVWSESDTSLAEGSVSCLDSESLGHNS-----NMGSDSGTMGSDSDE---ENV	296		
Db	222	ASRRQLNGRIYICPVHESATEMMVYLGTAAALNTNRYMINIRVDNIVQENSTPRLVIESV	281		
QY	297	AARASPEPELQRPYQMEVAQPALEGKNIICLPTSGKTRVAVYIAKHLDKKKASEP	356		
Db	282	RQRIHRQRQLCLRNQEEELCQVALQGNKTIVTAPTSGKTVIAAIIKEHFESRSSEGR	341		
QY	357	GKVIVLV-NKVLVLEQLFRKEFQPLKWKYRVIGLSGTQLKISFPEVVKSCDIISTAQ	415		
Db	342	FKALFMTPTNSMILNQ--AASISSYLDHYVHTQIIQSGD--NVPTRNVIQSKDLIVATPQ	397		
QY	416	ILENSLLNLENGEDAGVQ-----LSDFSLIIIDECHTNKEAVYNNIMRHYLMOKLKNR	470		
Db	398	MIVNLCNEHRNSLDDESRLDQFFLSTFTTIFFDECHNTVKNSPYSNIMREY--HYLKNMG	455		

Db 21 LESIYRDPKGEFCSELLPLGKIDELKTHS-ENSQEFKQLYQDLKSNLSIADP-----E 75
QY 129 LLVRDVL-----DKCMEBELLTIEDNRNRIAAENNGESGVRELLKRIQKENVFSAF 181
Db 76 RLYKDIMAYLOSFPKCIHLVDLVNYSGRN-VKLSFRYILDHLEGLRLIEPK-----VV 129
QY 182 LNVLRQTGN-NELVQELTSGDCSESNAEIEENLSQVDGQVVEBQLLSTTVQPNLEKEVWGM 240
Db 130 LNYLDSYPQYSEAVKLL--KDEIDRHEEDNN-----DPEFIKKLILRTVPLLGEQAAYDI 182
QY 241 ENNSSESPADSSVSVSESDTSLA-EGSVSCLDESIGHNSNMGSDSGTMGSDSDENVAAR 299
Db 183 MYTIWEKAQWTSIIINIRVDSIAPENSASRL-----VIESVRQR 221
QY 300 ASPEPELQRLPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKV 359
Db 222 IHQIQFCLRNQYQELCQVALQGNKTIIVTAPTGSCKTIVIAANIIEKHFESR---SSEG-- 276
QY 360 IVLVNKKVLLVEQLFRKEFPFLKKMYRVIGLSGDTQLKISFPEVVKSCDIIISQAILEN 419
Db 277 -----KRFKTQTIQSGD--NVPTRNVIQSKDLIVATPQIVN 311
QY 420 SLLNLENGEDAGVQ-----LSDPFLIIDECHHTNKBAVYNNIMRHYLMQKLKNNRLK 472
Db 312 -LCN-EHRDPLDDDEYPPPEQFFLSTFIIFDECHNTVKNSPYSNVMREY--HYLKNMGNM 367
QY 473 KENKPVIPLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQI 532
Db 368 PEGH---SPFIQIGLTASLTGDKKNCMQVRSYIAGLCANMDVKELSIKDNLELHDHN 424
QY 533 QEPCKKFAIADATREDP-----PKEKLEIMTRIQTQYQMSPMDFGTQP-----Y 578
Db 425 PFVTDQVSFCERSNDGPIEMFTKRLKQMMQVEVDLIRTLKNEPTVKYEIPPTDKEHNRY 484
QY 579 EQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKF 638
Db 485 ENWISQNRNCVSLAGSRNKTLLIEVL-----DVLKVQYMFQSKMIFWNYFRKYRP 534
QY 639 AV--IEDDSDEGGDDDEYCDGDEDDLKPLKLD-----ETDRFLMTLFFENNMMLK 688
Db 535 GFKRIFFQDQCFYALSYNINFNPEVALKKYLEKELGPERIRNFTDN--MNLWDN--CHR 590
QY 689 RLAENPEYENKLTKLNTIMEQVTRTEESARGIIFTKTRQSAAYALSQWITENEKFAEVG 748
Db 591 ELVGIGSAENPMIARTVQFILDQNEQTSDFRAIIFVTRKKEADFLNYVL--NDRLHELG 647
QY 749 VKAHLIG----AGHSSEFKPMTQNEQKEVISKFRGTGKINLLIATTVAEEGLDIKECNIVI 805
Db 648 IKSDWMSGQKSTASSADISASKQMEKLMKFADGENQILVSTSVAAEGLDIPESLVI 707
QY 806 RYGLVTNEIAMVQARGARADESTVVLVAHSGSGVIEH--ETVNDFREKMMYKAHCVQN 863
Db 708 KINYATNETAHVQRRGRARARNKCVLITNS---IALHVQESNNLAKENLMTETISLIQN 764
QY 864 MKPE-----EYAHKILELQMSIMEKKMKTKRNIKHYKNNPSLITFLCKNCSVLACSG 917
Db 765 SPGEFRQCVDEESNKVWPRIQREDTDKAQRIKEQI-----NRNIVYKIVCMKCDTVLCTN 819
QY 918 EDIHVIEKMHVNMTPF-----KELYIVRENKALQKKCADYQINGEIIIC---KCGQAWGT 970
Db 820 KDIR-SKNTQYIVCNPGFWSLVRRIPLEQRASNK-----FNSTGSIECLGERCSKLGQ 874
QY 971 MMVHKGLDLPCLKIRNFVVFKNSTKKQYKKWVEL-----PITFPNLD 1014
Db 875 LIDVNTVNLPCLVKKSILLIESTNERILVKQWKNIIDHEFTPTTLKQRD 924

RESULT 4
T46312

hypothetical protein DKFZp434J111.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C;Accession: T46312
R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23036

A;Accession: T46312

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-398 <AAA>

A;Cross-references: EMBL:AL137608

A;Experimental source: adult testis; clone DKFZp434J111.1

C;Genetics:

A;Note: DKFZp434J111.1

Query Match 9.7%; Score 516; DB 2; Length 398;

Best Local Similarity 33.3%; Pred. No. 2.2e-19;

Matches 141; Conservative 68; Mismatches 157; Indels 58; Gaps 14;

QY 594 NRKERV-----AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSDE 647

Db 2 DEESRICKALFLYTSHLRKYNDALIISEHARMKDALDYLDFFSNVRAAGF----- 52

QY 648 GGDDEYCDGDEDDLKPLKLDLTDTRFLMTLFFENNMMLKRLAENPEYENKLTKLNT 707

Db 53 -----DEIEQDL-----TQR-----FEELQELSVSRDPSNENPKLEDLCFI 90

QY 708 IMEQVTRTEESARGIIFTKTRQSAAYALSQWITENEK--FAEVGVKAHHLIGAGHSSEFKP 765

Db 91 LQEEYHLNPETIT-ILFVKTALVDALKNWIEGNPKLSFLKPGI----LTGRKTNQNTG 145

QY 766 MTQNEQKEVISKFR-TGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRAR 824

Db 146 MTLPAQKCIIDAFKASGDHNIATSVADEGIDIAQCNLVILYEVGVNVIKMTQTRGRGR 205

QY 825 ADESTVVLVAHSGSGVIEHETVNDFREKMMYKAHCVQNMKPEEYAHKILELQMSIMEK 884

Db 206 ARGSKCFLLT-SNAGVIEKEQINMYKEKMMNDSILRLQTWDEAVFREKILHIQTH---EK 261

QY 885 KMKTKRNIKHY--KNNPSTLITFLCKNCSVLACSGEDIHVIKMHVNMTPPEFKELYIVR 942

Db 262 FIRDSQEKPKVPDPDENKKL---LCRKCKALACYTADRVVIEECHYTVLGDAPKECFVSR 318

QY 943 ENKALQKKCADYQINGEIIIC---KCGQAWGTMMVHKGLDLPCLKIRNFVVFKNSTKKQ 999

Db 319 PHPK-PKQFSSEFKRAKIFCARQNCSDHWGIHVYKTFEIPVIKIESFVVEDIATGVQTL 377

QY 1000 YKKW 1003

Db 378 YSKW 381

RESULT 5

T20332

hypothetical protein D2005.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T20332

R;Wilkinson, J.

submitted to the EMBL Data Library, September 1996

A;Reference number: Z19258

A;Accession: T20332

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1134 <WIL>

A;Cross-references: EMBL:Z79752; PIDN:CAB02082.1; GSPDB:GN00019; CESP:D2005.5

A;Experimental source: clone D2005

C;Genetics:

A;Gene: CESP:D2005.5

A;Map position: 1

A;Introns: 33/2; 91/2; 116/3; 144/1; 319/1; 361/1; 421/3; 482/2; 591/1; 684/3; 757/1; 95

Query Match

Best Local Similarity 9.4%; Score 499; DB 2; Length 1134;

Matches 193; Conservative 134; Mismatches 288; Indels 206; Gaps 29;

QY 294 ENVAARASPEP-----ELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVY 341
 Db 345 EDVCTTIRPQKDSGYNPDVAVTELVRTYQEELVQPALEGKNCVIVAPTSGSKTEVAIY 404
 QY 342 IAKDHLDKKKAASEPGKVIIVLVNKKVQLFRKEFPFLKKYRVIGLSGDTQLKISFP 401
 Db 405 AALKHIEERTSQGKPSRVLLVPKIPLVGQQ-KDRFLKYCMGYEVNGFHG-SESSVSGT 462
 QY 402 ---EVKSCDIISTAQILENSLLENAGEDAGVQLSDFSLIIDECHHTNKEAVYNNIM 458
 Db 463 GRRDEVIATHSVMTPIILINMLQSVRQNE--RLYVSDFSMMIFDEVHKAANKHPY--VL 518
 QY 459 RHYLMQKLNKNNLKKENKPVIPLPQILGLTASPG--VGGATKQAKAEHILKLCANLDAF 516
 Db 519 INQMVEWQKYEK-----PQIIIGLTASLSVKVDGQXQDENQMLNDIYNMLALINAP 567
 QY 517 TIKTV--KENLDQKNOIQEPCKKFAIADATREDPFKE-----KLEIM----- 558
 Db 568 HLSTITROSSIDELNEHVGPDDSVLCLPAKENILRDYIERVYNHAGKFFLEELASMSK 627
 QY 559 -----TRIQYCOMSPMSDFGTQPYE-----QWAIQMEKKAAGKGNRKERVCAE 602
 Db 628 STGRNNTIPPNNMINTFKKNQPKN-----YFYDLSLLQIIQELNKLNVPEKWNSTWAK 681
 QY 603 HLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDSDSGDDDEYCDGDEDD 662
 Db 682 YMKVYLEARGIVDLMPAMVAFKYE-----KAIGKLNEHSE--TVEYSTFIKDHT 731
 QY 663 LKKPLKLDTRFLMTLFFENNKMKLRLAENPEYENKLTKLRLTIMEQYTRTEESARGI 722
 Db 732 LKQTIQSVPE-----IVLRL-----KQYT--HQSVPHQFGNYGEQMVGY 769
 QY 723 IFTKTRQSAYALSQWITENKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGK 782
 Db 770 VLGTNKQAV-----QTSQEQQTLTDKFNNGR 797
 QY 783 INLLIATTVAEEGLDIKECNIVIRYGLVT-NEIAMVQARGARADESTYVLVAHSGSGVI 841
 Db 798 LKVIVATSVVEEGLDVTACNLIIKYNCSGSAIQVQORGRARAKNSRVLLSVK-SSIN 856
 QY 842 EHETVNDFREKVMYKAIHCVQNMKPEEYAHKILELQMSIMEKMKTKRNIKHYKKNPS 901
 Db 857 ETETNALISEKYMRLCVKKITENGEKQLAEBVKRVAELNAAERKNLEEQNLRLRHENK 916
 QY 902 LITFLCKNCSVLACSGEDIHV-----IEKMHVNMTPEFKELYIVRENKALQ 948
 Db 917 IYKLMCSNCSKEFC--KSIYIKKVFNSNMYVDPDPSVWRFLHV-----ESVETFI---KCLK 966
 QY 949 --KKC--ADYQIN-----GEIIC-KCG 965
 Db 967 ITWKCRIDYQIABFPNFAFRQLTFRFLCNFQMFQKRKVSXYLSEDNQPLSDIKCFHCK 1026
 QY 966 QAWGTMVHKGLDPLCLKIRNFVVVFKNNSTKKQYKKWVEL 1006
 Db 1027 LDVGRAYKIRGYLPQLSVKALTFFVQESDYSSMTKAKWSDV 1067
 RESULT 6
 D75219
 ATP-dependent RNA helicase, eif-4a family PAB0190 - Pyrococcus abyssi (strain Orsay)
 C;Species: Pyrococcus abyssi
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 10-May-2001
 C;Accession: D75219
 R;anonymouse, Genoscope
 submitted to the EMBL Data Library, July 1999
 A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
 A;Reference number: A75001
 A;Accession: D75219
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-752 <KAW>
 A;Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49203.1; PID:g545771

A;Experimental source: strain Orsay
 C;Genetics:
 A;Gene: PAB0190
 C;Superfamily: ATP-dependent RNA helicase eif-4A
 F;33-40/Region: nucleotide-binding motif A (P-loop) #status atypical
 F;131-136/Region: nucleotide-binding motif B
 F;135-138/Region: DEAD/H motif
 Query Match 8.5%; Score 453; DB 2; Length 752;
 Best local similarity 25.3%; Pred. No. 9.4e-16;
 Matches 183; Conservative 124; Mismatches 259; Indels 156; Gaps 28;
 QY 306 LQLRFPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKAASEPGKVIIVLV-N 364
 Db 11 LQPRLYQ-ELIYAKCKERNCLVLPTGLGKTIIAMMIADYRLNK-----YGGKVLMLAPT 64
 QY 365 KVLIVE--QLFRKEFPFLKQYRVIGLSGDTQLKISPEVVKS---CDIIISTAQILEN 419
 Db 65 KPLVLQHAETFRKFFNLPEK---IVALTGE---VSPSERVKAWNRAKVIIATPQTVEN 117
 QY 420 SLLNLENGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMQKLNKLNKKNKPKVI 479
 Db 118 DLLT-----GKINLEDVVLVTFDEAHRAVGNVYVYIAREYLNQ-AKN----- 159
 QY 480 PLPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKF 539
 Db 160 --PHVIGLTASPG-----SSQEKILEVLRNLGIEHIEYRSENSPDVKPYVQG--IKF 207
 QY 540 AIADATREDPFKE--KLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNR-- 595
 Db 208 EWIKVELPGLYKEVRKLLRDMLE---DSLPLAEAGL--IESASADLPKKEILRAGQII 261
 QY 596 -KERVCAEH-LRK-----YNEALQINDTIRMIDAY-----THLETFTYNEEKDKKFAVIED 643
 Db 262 NEETAKGNHDLKLLLFHAMALKLHHAIELLETOGLSALRVYLKLYEEAK----- 312
 QY 644 DSDEGGDDEYCDGDEDEDDLLKPLKLDDETDRFLMTLFFENNKMKLRLAENPEYENKLT 703
 Db 313 -----TGSTRASKELFLDRMKKALALLIQ-----AKELGIDHFKMEV 350
 QY 704 LRNTIMEQYTRTEESARGIIFTKTRQSAYALSQWITENKFAEVGVKAHHLIGAGHSSEF 763
 Db 351 LKELVKEQLSK-KENSKIIIVFTNYRETARKVVEELTKE-----GIKAKRFVGOATKEND 403
 QY 764 KPMTQNEQKEVISKFRGTGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGR- 822
 Db 404 RGMQSREQKLILDSFARGEFNVLVATSVGEEGLDVPEVDLVVFYEPVPSAIRSVQRRGT 463
 QY 823 ARADESTYVLVAHSGSGVIEHETVNDFREKMYKAIHCV-----QNMKPEEYAHKILE 875
 Db 464 GRQKPGRVVILIAQGTREAYVSSKHKERQMRITIRVMSQAIRKREKQLSLESYVKKDE 523
 QY 876 LQMOSIMEKMKTKRNIKHYKKNPSLITFLCKNCSVLACSGEDIHVEIKMHVNMTPEF 935
 Db 524 KEMEREKEERKGEREGVKVVVDSRELRS-----EVVRLKTLGVKIEV 567
 QY 936 KEL-----YIVRENKALQKKCADYQING-----EIIICKGQAWGTM 972
 Db 568 KTLDVGDYIISDEVAIERKSANDFIQSIIDGRFLDQVKRLKDSYPRPVIVVEGQLYGI 627
 QY 973 VH 974
 Db 628 VH 629
 RESULT 7
 D71203
 probable ATP-dependent RNA helicase - Pyrococcus horikoshii
 C;Species: Pyrococcus horikoshii
 C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 10-May-2001
 C;Accession: D71203
 R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,

DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: D71203
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-650 <KAW>
A;Cross-references: GB:AP000007; NID:g3236134; PIDN:BAA31019.1; PID:g3258336
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1896
C;Superfamily: ATP-dependent RNA helicase eIF-4A
C;Keywords: ATP; nucleotide binding; P-loop
F;33-40/Region: nucleotide-binding motif A (P-loop)
F;131-136/Region: nucleotide-binding motif B
F;135-138/Region: DEAD/H motif

Query Match 8.3%; Score 441.5; DB 2; Length 650;
Best Local Similarity 25.9%; Pred. No. 3e-15;
Matches 185; Conservative 122; Mismatches 261; Indels 147; Gaps 29;
QY 306 LQLRPYQMEVAQPALEGKNIICLTGSGKTRVAVYIAKDHLDKKKKASEPGKVIVLV-N 364
Db 11 LQRLYQ-ELIYAKCKERNCLIVLPTGLGKTTIAMMIADYRLDK-----YGGKVLMLAPT 64
QY 365 KVLVVE--QLFRKEFQFLKWKYRVIGLSGDTQLKISFPEVVK-----CDIIISTAQILEN 419
Db 65 KPLVLQHAETFRKFFTLDPPEK--IVALTGE----ISPNERVRAWARAKVVIATPQTIIEN 117
QY 420 SLNLENGEDAGVQLSDFSLIIIDECHHTNKBAVYNNIMRHYLMQKLNRLKKNRPVI 479
Db 118 DLL-----VGRISLEDVTLIIFDEAHRAVGNVYAVYIAKEYLKQ-AKN----- 159
QY 480 PLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKF 539
Db 160 --PHVIGLTASPG-----STPEKIMEVLRLNGIEHIEYRSNSPDVKPYVQG--IKF 207
QY 540 AIADATREDPEKE--KLLIEMTRIQTCCMSDFGTQPYEQWAIQMEKKAAGKGNR-- 595
Db 208 EWVKVELPELYKEVRKLRDMLR---DSLKPLAEAGL--LDSTSADIPKKEILRAGQII 261
QY 596 -KERVCAEH-LRK-----YNEALQINDTIRMIDAY-----THLETFYNEEKDKKFAVIED 643
Db 262 NEEMAKGNHDLRLKLLFHAMALKLHHAIELETTQGLSALRVYLRLEYEAK----- 312
QY 644 DSDEGGDDEYCDGDEDDLLKPLKLDLTDRLMTLFFENKMLKPLAENPEYENEKLT 703
Db 313 -----AGSTRASKELFLDKRMKKALALIQ-----AKELGIDHPKMEV 350
QY 704 LENTIMEQYTRTEESARGIIFTKTRQSAVALSOWITENEKFAEVGVKAHHLIGAGHSSEF 763
Db 351 LKGLIREQLKR-KENSKIIVFTNYRETAKKV-----EELMKDGIKARRFVGQASREND 403
QY 764 KPMTQNEQKEVISKFRGKINLLIATTVAAEGLDIKECNIVIRYGLVTNEIAMVQARGR- 822
Db 404 RGMQKEQLILDAFARSEFNVLVATSVGEEGLDVPEVDLVIFYEPVPSAIRSVQRRGT 463
QY 823 ARADESTYVLVAHSGSGVIEHETVDFREKMMYKAHCVQNMKPEEYAHKILELQMSIM 882
Db 464 GRQKPGRVILMAQGTREDAEAYWSSRQKEKIMRETIRMVSQM-----VRKERQLSLESYV 518
QY 883 EKKMKTNRNIAKHYNKPNLSITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKEL---- 938
Db 519 KREVNEE---IKEEKETGGIKVWIDSRELRS-----EVKKLKTILGIKIEVRTLDVGD 569
QY 939 YIVRENKALOKKCADYQING-----EIIICKCGAWGTMVH 974
Db 570 YIVSEDAIERKSANDFTQSIIDGRLFDQVKRLKEAYPRPVIIIEGQLYGINRVH 624

RESULT 8
A69432

ATP-dependent RNA helicase homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 24-May-2001
C;Accession: A69432
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glocke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.L.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: A69432
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-741 <KLE>
A;Cross-references: GB:AE001002; GB:AE000782; NID:g2689325; PIDN:AAB89786.1; PID:g264910
C;Superfamily: ATP-dependent RNA helicase eIF-4A
C;Keywords: ATP; nucleotide binding; P-loop
F;36-43/Region: nucleotide-binding motif A (P-loop)
F;132-137/Region: nucleotide-binding motif B
F;136-139/Region: DEAD/H motif

Query Match 7.9%; Score 417; DB 1; Length 741;
Best Local Similarity 23.8%; Pred. No. 6.5e-14;
Matches 169; Conservative 114; Mismatches 230; Indels 198; Gaps 30;
QY 303 EPELQRPYQMEVAQPALEGKNIICLTGSGKTRVAVYIAKDHLDKKKKASEPGKVIVL 362
Db 11 ENTIERRMYYQISIAATALT-KNTLVVPTGLGKTTIAALVIASRL-----LNEDGKVLFL 64
QY 363 VNKVLLVEQLFRKEFQFLKWKYRV----IGLSGDTQLKISFPEVVKSCDIIISTAQILEN 419
Db 65 APTKPLVEQHAR-----FLKRVLKVEEIVSLSGEVPE-KRKELWEKARIVVSTPQVVEN 118
QY 420 SLNLENGEDAG-VQLSDFSLIIIDECHHTNKBAVYNNIMRHYLMQKLNRLKKNRPV 478
Db 119 DLL-----AGRISLEDVILVVFDEAHRAVGNVYAVYIAKEYLRTAKK----- 160
QY 479 IPLPQILGLTASPGVGATKQAKAEHILKLCANL--DAFTIKT----- 520
Db 161 ---PLILAMTASPG-----SDPERIMEVIQNLGIEAIEVRTEWDSVAPYVGKKRIE 209
QY 521 -----VKENLDQLKNQIQEPCCKFAIADATR-----EDPFKEKLLIEMTRIQTCCMS 569
Db 210 WIKVDIPEEMKEVKEKRLKE-CIKIRFKRLRELWIEVPENSSKRDLALQALQAEAAASSQ 268
QY 570 MSDFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQINDT--IRMIDAYTHLE 627
Db 269 SSEI----FEALSILAE-----IMKLOHAVELIETQGVKAVKSYLR-- 305
QY 628 TFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDDLLKPLKLDLTDRLMTLFFENKML 687
Db 306 -----KLIVREATSKGG-----SKAA 320
QY 688 KRLAENP-----EYENEKLTKLRLNTIMEQYTRTEESARGIIFTKTRQSAVALS 735
Db 321 KSIVGDPPIFKKAVIALSKCKVEHPKLEKLKEILKEQFEKNPDS-RVIVFTNYRDSAEML- 378
QY 736 QWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVSKFRGKINLLIATTVAAEEG 795
Db 379 --VNELSPLFPVA---KFVGQASRDNDKGMROKEQIETIDFRRGVYKVLVATSVGEEG 432
QY 796 LDIKECNIVIRYGLVTNEIAMVQARGR-ARADESTYVLVAHSGSGVIEHETVDFREKMM 854
Db 433 LDIPSTDLVVYEAVPSEIRAIQRKGTGRGREGRIVLVLTGKT-----RDEAY 481
QY 855 YKAHCVQNMKPEEYAH-KILELQMSIMEKMKTKRNIKAHYKNPNLSITFLCKNCSVL 913
Db 482 Y-----YSSMKKERKMYDKILEI--KRIDRK---QRSIGDYVLPEETGKIVIVDSREL 531
QY 914 ACSGEDIHVIEKMHVNMTPEFKEL-----YIVRENKALOKKCADYQINGEI 960

Db 12 PE-KIEARTYQQLAADVIRKGNMIVAPTALGKTVVAVLVAERLRYRGS-----KVLII 66

QY 362 L-VNKVLLVEQLFRKEFQFLKKWYRVIGLSGTQLKISF-PEVVK-----SCDIIISTAQ 415

Db 67 LSPSKPLAIOH--EESFEFM-----LATCTSLTSGIKPEERKERWIKSQIISATPQ 116

QY 416 ILENSLLNLENGEDAG-VQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKE 474

Db 117 TVESDIL-----AGRYDLRDVSLIVFDECHRAVGSYSYVFLASNY-----IQNAR----- 161

QY 475 NKPVPLPQILGLTASPGVGATKQAKAEHILKLCANL--DAFTIKT-----VKENLDQ 527

Db 162 -----HPLILGLTASPGAD-----EDKIKTVCCENLFMNEVVVKTEGDPDVRPYLKP 207

QY 528 LKNQIQEPCCKFAIADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEK 587

Db 208 IKIEWVK-----VRMTPELEDIRELLRKLKLN-RLKMLKNLGLVID-----TISVGK 252

QY 588 KAAKKG-----NRKERVCAEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIE 642

Db 253 KDLKARGRVQNRARSTSPPRACYRAISLLASCINVEHALELET----- 298

QY 643 DSDDEGGDDEYCDGEDDDLKPLKLDDETDRFLMTLFFENNKMRLKLAENPEY----- 696

Db 299 ----Q3-----IRPL-----HQVLLRLKEKTKAAKGLLADPDFTRAMEHL 334

QY 697 -----ENEKLTCLRNTIMEQYTRTESARGIIFTKTRQSAVALSQWITENEKFAEV 747

Db 335 TRRAMMSGVEHPKLDRLME-ILKRELKGD-ARIIVFTQFRDTEEIYQ-----RCKRE 386

QY 748 GVAHHHLIGAGHSSEFKMTQNEQKEVSKFRGKINLLIATTVAEEGLDIKECNIVIRY 807

Db 387 GINAVKFGQNSRSGEGLTKQQRDIKSFMRGNHVDVLLSTVAEEGIDIPSVDLVVMY 446

QY 808 GLVTNEIAMVQARGARADESTYVLVAHSGSGVTEHETVNDFREKMYKAIHCVQNMKPE 867

Db 447 EPVPSEIRMIQRRGRTGRKRGRMVV-----LITEKTRD---EAYYSSIRKERSMK-E 496

QY 868 EYAHKILELQMQSIMKKMKTKRNIKHYKNPNSLITFLCKNCSVLACSGEDIHVIEKMH 927

Db 497 NLRGGSVNVEVNPIMEP-----SGEGPFIYADSR 525

QY 928 HVN--MTPEFKELYIVRENKALQKKCADYQINGELI 961

Db 526 EVNSRVRLKKGIVDFELKPL--AVGDYQISEDTI 559

RESULT 13

D84386

ATP-dependent RNA helicase homolog eIF-4A [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 10-May-2001

C;Accession: D84386

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li

A;Title: Genome sequence of Halobacterium species NRC-1.

A;Reference number: A84160; MUID:20504483; PMID:11016950

A;Accession: D84386

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-784 <STO>

A;Cross-references: GB:AE004437; NID:gl0581766; PIDN:AAG20456.1; GSPDB:GN00138

C;Genetics:

A;Gene: eif4a

C;Superfamily: ATP-dependent RNA helicase eIF-4A

Query Match 5.9%; Score 313.5; DB 2; Length 784;

Best Local Similarity 21.6%; Pred. No. 1.5e-08;

Matches 126; Conservative 91; Mismatches 190; Indels 175; Gaps 22;

QY 306 LQLRPYQMEVAQPALEGKNIICLPTSGKTRVAVYIAKDLHDKKKKASEPGKVIVLVNK 365

Db 18 IEARQYQLQAAAAAQG-HTLVCLPTGLGKTTVLLVTAYRL-----ADDAGGTALLAPT 72

QY 366 VLLVEQLFRKEFPFLKKWYR-----VIGLSGTQLKISFPEVVK-----CDIIS 412

Db 73 KPLVEQ-----HAGFYREALAIPDDDVVVFGETR-----PDRRAAWTDARVVVA 118

QY 413 TAOILENSLLNLENGEDAGVQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLK 472

Db 119 TPQVVENDLVG-----GRIDMDVVHCTFDECHRTATGDYATYVAERYHAD----- 164

QY 473 KENKPVIPLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQI 532

Db 165 -----AAAPLVTAMSASPG-----GTEAEIRTVCCENLGVGNVEVMTEDE----- 202

QY 533 QEPCKKFAIADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYE--QWAIQMEKKA 589

Db 203 -----DA-----DVGEHTHDTDVQW----- 217

QY 590 AKKGNRKERVCAEHLRKYNEALQINDTIRMI--DAYTHLET-----YNEE 633

Db 218 -----ERVTLF-----EILEVRDAINDVIEDRLAKLREIGVTKASSPDISQKDLNEI 265

QY 634 KDKKFAVIEDSDDEGGDDEYCDGEDDDLKPLKLDDETDRF-LMTLFFEN----- 683

Db 266 RARLQQLIDDDSDGYQGSVHAENVK--LKRAVELVETQSVEVRRYFERQORNAANTAG 323

QY 684 -NKMLKRLAENPEYENEKLT-----KLRNT--IMEQYTRTESARGIIFTKTRQSA 731

Db 324 ASKASQRLVSEPAVKRAMRTAREFDGLHPKFRQARMLLAETLGLIEDGDRVIVFTESRDTA 383

QY 732 YALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMQNEQKEVSKFRTGKINLLIATTV 791

Db 384 EALTAFLEGEH-----FDTRRFVGQGDADGSDGMTQTEQRETLAEFRNGDFEVLVSTSV 436

QY 792 AEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLV 833

Db 437 AEEGLDVPEVDLVLFPEPPTAIRSVQRKGRGTGRQTAGRVVV 478

RESULT 14

S48436

probable RNA helicase YIR002c - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein YIR2c

C;Species: Saccharomyces cerevisiae

C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Apr-2002

C;Accession: S48436; S50885

R;Badcock, K.; Churcher, C.

submitted to the EMBL Data Library, August 1994

A;Reference number: S48432

A;Accession: S48436

A;Molecule type: DNA

A;Residues: 1-993 <BAD>

A;Cross-references: GB:Z47047; EMBL:Z38062; NID:g603997; PID:g763347; MIPS:YIR002c

R;Voss, H.; Tamames, J.; Teodoru, C.; Valencia, A.; Sensen, C.; Wiemann, S.; Schwager, C.

Yeast 11, 61-78, 1995

A;Title: Nucleotide sequence and analysis of the centromeric region of yeast chromosome 1

A;Reference number: S50795; MUID:95282515; PMID:7762303

A;Accession: S50885

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-993 <VOS>

A;Cross-references: EMBL:X79743

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994

C;Genetics:

A;Gene: SGD:MPH1

A;Cross-references: SGD:S0001441

A;Map position: 9R

C;Keywords: ATP; nucleotide binding; P-loop

F;107-114/Region: nucleotide-binding motif A (P-loop)

F;205-210/Region: nucleotide-binding motif B

F;209-212/Region: DEAD/H motif

Query Match
Best Local Similarity 20.3%; Pred. No. 6.7e-08;
Matches 143; Conservative 118; Mismatches 246; Indels 197; Gaps 25;

QY 307 QLRPYQMEVAQPALGKNIICLPTSGSKTRVAVYIAKDHLDKKKASEPGKVIVLVNKV 366
Db :|||:::||:|||||::|||::|||:
86 EVRDYQYTIVHKSL-FQNTLCIPTGMKTFTASTVMNIFRWTKKA----KIIFTAPTR 140

QY 367 LLVEQLFRKEFQFLKKWYRVIGLSGD-----TQLKISFPPEVKSCDIIISTAQILENS 420
Db |||::||:|||||::|||::|||::|||:
141 PLVAQQ-----IKACLGITGPSDQTALLDDKSRRNEEIWANKRVEFFATPQVVEND 192

QY 421 LNLNGEDAGVQLSDFSLIIDECHHTNKEAYNNIMRHLYMQKLKNNRLKKENKPVP 480
Db |::|::|||::|||::|||::|||::|||:
193 L-----KRGVLDPKDIVCLVIDEAHRATGSSAYTNVVKF-----IDRFNSSY-- 234

QY 481 LPQILGLTASPGVGATKQAEEHILKLANLDRAFTIK-TVKENLDQLKNQIQEPCKKF 539
Db ::|||::|::|||::|||::|||::|||::|||:
235 --RLALTATP-----ASDLGVQEVDVNLDISKIEIRTEESMDIVKYMKRKKEKI 284

QY 540 AIADATREDPFKEKLEIMTRIOTYCQMSPMSDFGTQPYEQWAIQM-----EKAA 590
Db |||||::|::|||::|||::|||::|||::|||:
285 EV-----PLLEIEDIE-----QLGMVAKPVLOAQAIELGIYECDPSQINAF 327

QY 591 KGNRKERVCA---EHLRKYNEAL---QINDTIRMIDAYTHLETPEYNEEKDKKFAV 640
Db |:|::|::|||::|||::|||::|||::|||:
328 KAMQSQKIIANPTPIPEGIKWRNFILQLLNNVGOMLKRLKIY-GIRTFNYFNK---- 382

QY 641 IEDDSDEGGDBEYDCGDEDLKKPLKLDETDRFLMTLFFEN-NKMUKLAEN----PE 695
Db -:-:-:-CTEFTTKYNLKK-----STNKIAAEFYHYHPILKNIKQCENYLSDPK 424

QY 696 YENE-KLTKLNTIMEQYTRTESARGIIFTKTROSAVALSQWI----- 738
Db |||::||:|||||::|||::|||::|||:
425 FVGHGKLCQVEDLMDFFOKRGSDSRVIIFTELRESALEIVKFIDS VADDQIRPHIFIQ 484

QY 739 -----TENEFKPAEVG-----VKAHHLLIGAGH 759
Db ARAKEGFDEVKYTRKHAPKGGRKKVERLHQEQEKFLEAERTKRAANDKLER SARRTGSSE 544

QY 760 SSEFKPMTONEQKEVISKFGTGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVOA 819
Db ::|||::|::|||::|||::|||::|||::|||:
545 EAQISGMNQMKQKEVIHFNFKKGEYNVLVCTSIGEGLDIGEVDLIICYDTTSSPIKNIQR 604

QY 820 RGR-ARADESTYVLVAHSVSGSVIEHETVNDPREKKMYKAHCQVNMK-----PE- 867
Db |||::||:|||||::|||::|||::|||:
605 MGRTGKRKDGKIVLLFSSNESYKFERAMEDYSTLQALISKQCIDYKXSDRIIPEDIPEC 664

QY 868 -----EYAHKILELQMQSIMKKMKTNRNIAK 894
Db HETLITINDENEINIEMEDVDEVIRYATOCMMGKKVKKPKKAITK 708

RESULT 15
T31323
ATP-dependent RNA helicase, eIF-4A family - Cenarchaeum symbiosum
C:Species: Cenarchaeum symbiosum
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T31323
R:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.
J. Bacteriol. 180, 5003-5009, 1998
A:Title: Genomic analysis reveals chromosomal variation in natural populations of the ur
A:Reference number: Z20994; MUID:98422450; PMID:9748430
A:Accession: T31323
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-502 <SCH>
A:Cross-references: EMBL:AF083072; NID:g3599393; PID:g3599409; PIDN:AAC62714.1

Query Match 5.3%; Score 284; DB 2; Length 502;
Best Local Similarity 24.3%; Pred. No. 2.6e-07;
Matches 133; Conservative 94; Mismatches 207; Indels 114; Gaps 23;

QY	303	EP-ELQLRPYQMEVAQPALEGKNIICLPTSGKTRVAVYIAKDHLDKKKKASEPGKVIV	361
Db	12	EPGAVERDDYQVGLAEQAIR-ENCIVVLPTGLGKTAVALQVIAHYLDEGRGALFLAPTRV	70
QY	362	LVNKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQL--KISFPEVVKSC--DIIISTAQIL	417
Db	71	LVN-----QHRQFLG---RALTISDITLVTGEDITPRRKKAWGGSVICATPEIA	116
QY	418	ENSLNLNLENGEDAGVQLSDFSLIIDECHHTNKEAVVNNIMRHYLMQKLKNNRLKKNKP	477
Db	117	RNDI-----ERGLVPLEQFGLVIFDEAHRAVGDYAYSSIAR-----A	153
QY	478	VIPLPQILGLTASPGVGATKQAKAAEEHILKLCANLDAPTIKTVKENLDQLKNQIQBPCK	537
Db	154	VGDNSRMVGMTATL---PSEREKADE---IMGTLLSRISIAQRTEDDPVKPYQETAT	205
QY	538	KFAIADATREDPFKEKLEIMTRIQTQYCOMSPMS-DEGTQPYEQWAIQMEKKAACKGNPK	596
Db	206	BWIKVDLPPEMKEIQRLKLALD-ERYSSLKRCGYDLGSRNLSALLRL-RMVVLGNNRR	263
QY	597	ERVCAEHLRKYNEALQINDTIRMIDAY--THLETFFNEEKDKKFAVIEDDSDEGGDDEYC	654
Db	264	A-----AKPLTAIRITYALNIFEAHGVTPFLKPCERTSKKKGVGVAE-----	306
QY	655	DGDEDEDDLKPLKLDETDRFLMTLFFENNMKMLKRLAENPEYENEKLTCLRNTIMEQYTR	714
Db	307	-----LFEQDR-----NFTGAIRAKAAQAAGMEHPKIPKLEDAV-----	341
QY	715	TEESARG--IIFTKTRQSAVALSQWITENEKFAEYGVKHAHLIG-AGHSSEFKPMTQNEQ	771
Db	342	--RGARGKALVFTSYRDSVDLI-----HSRLKAAGINSILIGKAGE---KGLKQRKQ	389
QY	772	KEVSKFRTGKINLLIATVAEEGLDIKECNIVIRYGLVTNEIAMVQARGR-ARADESTY	830
Db	390	VETVAKFRDGGYDVLVSTRVCEEGLDISEVNLVIFYDNPSSIRYVQRRGRTRGRKDAGRL	449
QY	831	VLVAHSGS	838
Db	450	IVLMAKGT	457

Search completed: June 2, 2004, 19:12:27
Job time : 31 secs

RESULT 15
T31323
ATP-dependent RNA helicase, eIF-4A family - Cenarchaeum symbiosum
C:Species: Cenarchaeum symbiosum
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T31323
R:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.
J. Bacteriol. 180, 5003-5009, 1998
A:Title: Genomic analysis reveals chromosomal variation in natural populations of the un
A:Reference number: Z20994; MUID:98422450; PMID:9748430
A:Accession: T31323
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-502 <SCH>
A:Cross-references: EMBL:AF083072; NID:g3599393; PID:g3599409; PIDN:AAC62714.1

Query Match 5.3%; Score 284; DB 2; Length 502;
Best Local Similarity 24.3%; Pred. No. 2.6e-07;
Matches 133; Conservative 94; Mismatches 207; Indels 114; Gaps 23;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:01:12 ; Search time 18 Seconds
(without alignments)
2965.106 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSNGYSTDENFRYLISCFRA.....LPITFPNLDYSECLFSDSD 1025

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	406	7.6	1909	1 DICE ARATH	Q9sp32 arabidopsis
2	402.5	7.6	1374	1 YQ3D_SCHPO	Q09884 schizosacch
3	392.5	7.4	778	1 YF05_METJA	Q58900 methanococc
4	357	6.7	1912	1 DICE_HUMAN	Q9upy3 homo sapien
5	354	6.7	1906	1 DICE_MOUSE	Q8r418 mus musculu
6	350	6.6	1845	1 DCRI_CAEEL	P34529 caenorhabdi
7	303.5	5.7	993	1 YIS2_YEAST	P40562 saccharomyc
8	223	4.2	557	1 YQHH_BACSU	P54509 bacillus su
9	217.5	4.1	2663	1 CENE_HUMAN	Q02224 homo sapien
10	214.5	4.0	1790	1 USO1_YEAST	P25386 saccharomyc
11	205.5	3.9	1583	1 GCC2_HUMAN	Q81wJ2 homo sapien
12	204.5	3.9	2163	1 BRR2_YEAST	P32639 saccharomyc
13	197.5	3.7	1163	1 SBCC_CLOAB	Q97fk1 clostridium
14	197	3.7	1679	1 GCC2_MOUSE	Q8cbg3 mus musculu
15	197	3.7	2230	1 GOA4_HUMAN	Q13439 homo sapien
16	195	3.7	3259	1 GOB1_HUMAN	Q14789 homo sapien
17	193.5	3.6	663	1 UVRB_STAAM	Q99vl7 staphylococ
18	193.5	3.6	663	1 UVRB_STAAM	Q8nzm0 staphylococ
19	193.5	3.6	2017	1 MYSN_DROME	Q99323 drosophila
20	193	3.6	715	1 HELS_PYRAB	Q9v0a9 pyrococcus
21	192.5	3.6	1357	1 KTN1_HUMAN	Q86up2 homo sapien
22	191	3.6	3911	1 AKA9_HUMAN	Q99996 h a-kinase
23	189.5	3.6	3210	1 CENF_HUMAN	P49454 homo sapien
24	188.5	3.5	1433	1 REST_CHICK	Q42184 gallus gall
25	187.5	3.5	720	1 HELS_PYRFU	Q73946 pyrococcus
26	185	3.5	2611	1 BF1E_MOUSE	Q91zn8 mus musculu
27	183.5	3.5	663	1 UVRB_FUSNN	Q8rgr2 fusobacteri
28	183.5	3.5	1875	1 MLP1_YEAST	Q02455 saccharomyc
29	182.5	3.4	3660	1 DMD_CHICK	P11533 gallus gall
30	181.5	3.4	2869	1 RBP1_PLAVB	Q00798 plasmodium
31	180.5	3.4	479	1 DBPA_BACSU	P42305 bacillus su
32	180.5	3.4	976	1 SCP1_HUMAN	Q15431 homo sapien
33	180.5	3.4	1131	1 YAB9_YEAST	P31380 saccharomyc

34	180	3.4	796	1 DD27_HUMAN	Q969q7 homo sapien
35	180	3.4	1690	1 C190_DROME	Q9vje5 drosophila
36	180	3.4	3214	1 BPA1_HUMAN	Q03001 homo sapien
37	179.5	3.4	607	1 DB10_NICSY	P45942 nicotiana s
38	179.5	3.4	1938	1 MYH6_MOUSE	Q02566 mus musculu
39	179.5	3.4	2022	1 ANTI1_ONCVO	P21249 onchocerca
40	179	3.4	564	1 ROK1_YEAST	P45818 saccharomyc
41	179	3.4	662	1 UVRB_STRPN	Q54986 streptococc
42	179	3.4	662	1 UVRB_STRR6	Q8dpk7 streptococc
43	179	3.4	754	1 YAJ3_SCHPO	Q09903 schizosacch
44	179	3.4	2748	1 NUM1_YEAST	Q00402 saccharomyc
45	178.5	3.4	661	1 UVRB_STAEP	Q8cpz0 staphylococ

ALIGNMENTS

RESULT 1
DICE ARATH
ID DICE ARATH STANDARD; PRT; 1909 AA.
AC Q9SP32; Q9FDY6; Q9MAN0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Endoribonuclease Dicer homolog (EC 3.1.26.-) (CARPEL FACTORY protein)
DE (SHORT INTEGRANTS 1 protein) (SUSPENSOR1 protein).
GN CAF OR SIN1 OR SUS1 OR ATIG01040 OR T25K16.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND MUTANT CAF-1.
RX STRAIN=cv. Wassilewskija;
RX MEDLINE=20025923; PubMed=10556049;
RA Jacobsen S.E., Running M.P., Meyerowitz E.M.;
RT "Disruption of an RNA helicase/RNase III gene in Arabidopsis causes
RT unregulated cell division in floral meristems.";
RL Development 126:5231-5243(1999).
RN [2]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, AND
RP MUTANTS SIN1-1 AND SIN1-2.
RC STRAIN=cv. Landsberg erecta;
RX MEDLINE=22364007; PubMed=12376646;
RA Golden T.A., Schauer S.E., Lang J.D., Pien S., Mushegian A.R.,
RA Grossniklaus U., Meinke D.W., Ray A.;
RT "SHORT INTEGRANTS1/SUSPENSOR1/CARPEL FACTORY, a Dicer homolog, is a
RT maternal effect gene required for embryo development in Arabidopsis.";
RL Plant Physiol. 130:808-822(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Paterson J., Pham P.K., Rizzo J.M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana";
RL Nature 408:816-820(2000).

DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative helicase C188.13c (EC 3.6.1.-).
GN SPC188.13c OR SPC584.10C.

OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,

RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe.";

RL Nature 415:871-880(2002).

CC -!- SIMILARITY: Belongs to the helicase family.

CC -!- SIMILARITY: Contains 2 RNase III domains.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AL032824; CAB37423.1; -

DR EMBL; AL049662; CAB41233.1; -

DR PIR; T39130; S62524.

DR GeneDB_SPombe; SPC188.13c; -

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR005034; DUF283.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR000999; RNase_III.

DR Pfam; PF00270; DEAD; 1.

DR Pfam; PF03368; DUF283; 1.

DR Pfam; PF00271; helicase_C; 1.

DR Pfam; PF00636; Ribonuclease_3; 2.

DR SMART; SM00487; DEXDc; 1.

DR SMART; SM00490; HELICc; 1.

DR SMART; SM00535; RIBOc; 2.

DR PROSITE; PS00517; RNase_3_1; 1.

DR PROSITE; PS0142; RNase_3_2; 2.

KW Hypothetical protein; Helicase; ATP-binding; Hydrolase; Nuclease;

KW Endonuclease; Repeat.

FT NP_BIND 32 39 ATP (POTENTIAL).

FT SITE 145 148 DECH BOX.

FT DOMAIN 916 1038 RNase III 1.

FT DOMAIN 1083 1233 RNase III 2.

SQ SEQUENCE 1374 AA; 158039 MW; 89AE9EF8DE7966C6 CRC64;

Query Match 7.6%; Score 402.5; DB 1; Length 1374;

Best Local Similarity 25.6%; Pred. No. 2.1e-12;

Matches 175; Conservative 107; Mismatches 244; Indels 157; Gaps 32;

QY 304 PELQLRPYQMEVAQPALEGKNVILCLPTGSGKTRVAVYIAKHLDKK-----KKA 353

Db 9 PQL-LRKYQDDVYNTASK-QNTLLVMRTGAGTKLLAVKLIKQKLEEQILIQESNLEHKKI 66

QY 354 SEPGKIVLVNKKVLLVEQ--LFRKEFQPFLLKKWYRVIGLSGDTQLKISPFVVKSCDII 410

Db 67 S-----VFLVNVKVPVLFQAEYIRSQLPAKVGMYFYGELSIEMLQ-----LTNIIKYNVI 118

QY 411 ISTAQILENSLLNLENGEDAGVQLSDFSLIIIDCHHTNKEAVYNNIMRHVLMQKLKNR 470

Db 119 VITADLF--YLFARG--FLSINDLNLIIFDECHHAIGNDAYARIMNDFY-----HR 166

QY 471 LKKE-NKPVIPPLQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLK 529

Db 167 AKAVLSKKHFTLPRIFGWTASP-----FTGK--KGNLYHRL 200

QY 530 NQIQEPCCKFAIADATREDPFKEKLEIMTRIQTQYQMSPMSDFGTQYEQWAIQMEKKA 589

Db 201 YQ-----WEQLFDSKAHV--VSENELADYFCLPESYVVMYNSKLV 238

QY 590 AKKGNRKERVCAEHL-----RKYNEALQINDTIRM-----IDAYTHLETFYNEEKDKK 637

Db 239 VPPSDSIKKCEETLQCKLISRAVKTALA--ETIDMGLWFGQVWLVYDFVETKRLKK 296

QY 638 FAVIEDDSDEGGDDDEYCDGDEDDLKPLKLDDETRFLMTLFFENNMKMLKRLAENPEYE 697

Db 297 KALGKQLSD-----DEE-----LAIDRLKIFVED--WKNNKYSNDGPRIPVFD 337

QY 698 N-----EKLTKLRNTIMEQYTRTEESARGIIFTKTRQSAYALSQWITENEKFAEVGVKAHH 753

Db 338 STDVTDKVKFLLLELLKATY-RKSDSVRTVIFVERKATAFTLSLFM---KTLNLPNIRAH 393

QY 754 LIGAGHS--SEFKPMTQNEQKEVSKFRGTGKINLLIATTVAEGLDIKECNIVIRYGLVT 811

Db 394 FIGHGPSDQGEFS-MTFRRKQKDTLHKFKTKYNVLIATAVAEEGIDVPSCNLVIRFNICR 452

QY 812 NEIAMVQARGRARESTYVLVAHSGGVIE-----HETVN-----DFREKMYKAIHCV 861

Db 453 TVTQYVQSRGRARAMASKFLIFLNTTELLIHERILHEEKLFALSELSNSNIFDSLVC 512

QY 862 QNMK-PEEYAHKILE-----LQMQSIMEKMKMT-KRNIKHY-----KNNPSLITFLC- 907

Db 513 EREVTDDIVVEVGETGALLTGLYAVSLLYNFCNTLSRDVYTRYPTTAQPCLSGWYCF 572

QY 908 -----KNCSVLACSGEDIHVIEK 925

Db 573 EVELPKACKVPAAGSPAKSIRK 595

RESULT 3

YF05_METJA

ID YF05_METJA STANDARD; PRT; 778 AA.

AC Q58900;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Putative ATP-dependent RNA helicase MJ1505.

GN MJ1505.

OS Methanococcus jannaschii.

OC Archaea; Euryarchaeota; Methanococci; Methanococcales;

OC Methanocaldococcaceae; Methanocaldococcus.

OX NCBI_TaxID=2190;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;

RX MEDLINE=96337999; PubMed=8688087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,

RL DNA Res. 9:99-106(2002).

RN [5]

RP SEQUENCE OF 1238-1912 FROM N.A.

RC TISSUE=Lung;

RX MEDLINE=99162526; PubMed=10051563;

RA Provost P., Samuelsson B., Radmark O.;

RT "Interaction of 5-lipoxygenase with cellular proteins.";

RL Proc. Natl. Acad. Sci. U.S.A. 96:1881-1885(1999).

CC -!- FUNCTION: Involved in cleaving double-stranded RNA in the RNA

CC interference (RNAi) pathway. It produces 21 to 23 bp dsRNAs

CC (siRNAs) which target the selective destruction of homologous

CC RNAs.

CC -!- SIMILARITY: Belongs to the helicase family.

CC -!- SIMILARITY: Contains 2 RNase III domains.

CC -!- SIMILARITY: Contains 1 PAZ domain.

CC -!- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AB028449; BAA78691.1; ALT_INIT.

DR EMBL; AJ132261; CAB38857.2; -.

DR EMBL; AB023145; BAA76772.2; ALT_INIT.

DR Genew; HGNC:17098; DICER1.

DR MIM; 606241; -.

DR GO; GO:0005622; C:intracellular; NAS.

DR GO; GO:0003725; F:double-stranded RNA binding; IDA.

DR GO; GO:0004525; F:ribonuclease III activity; IDA.

DR GO; GO:0030423; P:RNA interference, targeting of mRNA for des. . .; IEP.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001159; DS_RBD.

DR InterPro; IPR005034; DUF283.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR003100; PAZ.

DR InterPro; IPR000999; RNase_III.

DR Pfam; PF00270; DEAD; 1.

DR Pfam; PF00035; dsrm; 1.

DR Pfam; PF03368; DUF283; 1.

DR Pfam; PF00271; helicase_C; 1.

DR Pfam; PF02170; PAZ; 1.

DR Pfam; PF00636; Ribonuclease_3; 2.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00358; DSRM; 1.

DR SMART; SM00490; HELICC; 1.

DR SMART; SM00535; RIBOC; 2.

DR PROSITE; PS0137; DS_RBD; 1.

DR PROSITE; PS00821; PAZ; 1.

DR PROSITE; PS00517; RNase_3_1; 1.

DR PROSITE; PS0142; RNase_3_2; 2.

KW Helicase; ATP-binding; RNA-binding; Hydrolase; Nuclease; Endonuclease;

KW Repeat.

FT NP_BIND 34 41 ATP (POTENTIAL).

FT SITE 165 168 DECH BOX.

FT DOMAIN 881 1032 PAZ.

FT DOMAIN 1266 1393 RNase III 1.

FT DOMAIN 1656 1814 RNase III 2.

FT DOMAIN 1839 1904 DRBM.

FT CONFLICT 65 80 VLLTKLSYQIRGDFS -> STTLKSLYLDGETSA

FT CONFLICT 179 179 I -> F (IN REF. 1).

FT CONFLICT 185 185 N -> I (IN REF. 1).

FT CONFLICT 204 204 C -> W (IN REF. 1).

FT CONFLICT 208 208 E -> D (IN REF. 1).

FT CONFLICT 213 213 I -> F (IN REF. 1).

FT CONFLICT 383 384 QQ -> HS (IN REF. 1).

FT CONFLICT 482 483 KQ -> NT (IN REF. 1).

FT CONFLICT 599 599 D -> H (IN REF. 1).

SQ SEQUENCE 1912 AA; 217627 MW; 996399DB4B074F21 CRC64;

Query Match 6.7%; Score 357; DB 1; Length 1912;

Best Local Similarity 24.6%; Pred.No. 5.9e-10;

Matches 160; Conservative 88; Mismatches 203; Indels 200; Gaps 29;

QY 309 RPYQMEVAQPALEGKNIICLP TSGKTRVAVYIAKD-HLDKKKASEPGK-VIVLVNKV 366

DB 35 RKYQVELLEAALD-HNTIVCLNTSGKTFIAVLLTKELSYQIRGDFSRNGKRTVFLVNSA 93

QY 367 LLVEQLFRKEFQPFELKKWYRVIGLSGDTQLKI---SPFEVVKSC-----DII 411

DB 94 NQVAQ-----QVSAVRTHSDLKVGYSNLEVNASWTKERNQEFTHQVLI 139

QY 412 STAQILENSLNLNLENGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMQKLKNNRL 471

DB 140 MTCYVALNVLKN-----GYLSLSDINLLVDFDECHLAILDHPYREIMK----- 181

QY 472 KKENKPVIPLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQ---- 527

DB 182 LCENCP--SCPRIILGTASI-LNGKCDPEELEKIQKL-----EKILKSNAETATDL 230

QY 528 --LKNQIQEPCKKFAIAD---ATREDPFKEKLLEI-----MTRIQTTC 565

DB 231 VLDRYTSQPC--IVVDCGPFTRSGLYERLLMELEALNFINDCNISVHSKERDSTLI 288

QY 566 QMSPMDS-----FGTQPYEQWAIQMEKKA--KGNRKE-----RV 599

DB 289 SKQILSDCRAVLVGLPWCADKVGMMVRELQKVIKHEQEELHRRKLLFTDTFLRKIHAI 348

QY 600 CAEHLRKYN-----EALQINDTIRMIDAYTHLE---TFYNEEKDKKFAVIEDDSDE 647

DB 349 CEEHFSPASLDLKFVTPKVIKLEILRKYPYERQQFESVEWYNNRNQDNYSWSDSED- 407

QY 648 GGDDEYCDGDEDDLKPKLKDDETRFLMTLFFENNMKMLKRLAENPEYENEKLTCLRNT 707

DB 408 --DDE---DEEIEKEKP-----ETNFPSPFTNI 431

QY 708 IMEQYTRTEESARGIIFTKTROSAVALSQWITENEK-----FAEVGVKAHHLIGAGH-- 759

DB 432 LC-----GLIFVERRYTAVVLNRLIKEAGKQDPDELAYISSNFTIGHGIGKNQPR 480

QY 760 ----SSEFKPMTQNEQKEVISKFTGKINLLIATTVAEGLDIKECNIVIRYGLVTVNEIA 815

DB 481 NKQMEAEFR-----KQEEVLKRFRAHETNLLIATSIVEEGVDIPKCNLVVRFDLPT 535

QY 816 MVQARGRARADESTYLVVAHSGSGVIEHETVNDFREKM-MYKAHCVQNMK 865

DB 536 YVQSKGRARAPISNYIMLA-----DIDKIKSFEEDLTKYKAIEKILRNK 579

RESULT 5

DICE MOUSE

ID DICE MOUSE STANDARD; PRT; 1906 AA.

AC Q8R4I8;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Endoribonuclease Dicer (EC 3.1.26.-) (Double-strand-specific

DE ribonuclease mDCR-1).

GN DICER1 OR DICER OR MDCR.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=21886641; PubMed=11889553;

RA Nicholson R.H., Nicholson A.W.;

RT "Molecular characterization of a mouse cDNA encoding Dicer, a

RT ribonuclease III ortholog involved in RNA interference.";

RL Mamm. Genome 13:67-73(2002).

RN [2]

SEQUENCE FROM N.A.

OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=95282515; PubMed=7762303;
RA Voss H., Tamames J., Teodoru C., Valencia A., Sensen C., Wiemann S.,
RA Schwager C., Zimmermann J., Sander C., Ansoerge W.;
RT "Nucleotide sequence and analysis of the centromeric region of yeast
RT chromosome IX";
RL Yeast 11:61-78(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313266; PubMed=9169870;
RA Churche C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moulle S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX";
RL Nature 387:84-87(1997).
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DEAH
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X79743; -; NOT ANNOTATED_CDS.
DR EMBL; Z38062; CAA86204.1; -.
DR PIR; S48436; S48436.
DR Germline; 139714; -.
DR SGD; S0001441; MPH1.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0003724; F:RNA helicase activity; IMP.
DR GO; GO:0006281; P:DNA repair; IMP.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002464; DEAH_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00690; DEAH_ATP_HELICASE; FALSE NEG.
KW Hypothetical protein; ATP-binding; RNA-binding; Helicase.
FT NP_BIND 107 114 ATP (POTENTIAL).
FT SITE 209 212 DEAH BOX.
SQ SEQUENCE 993 AA; 114057 MW; 474DDC99C543171F CRC64;
Query Match 5.7%; Score 303.5; DB 1; Length 993;
Best Local Similarity 20.3%; Pred. No. 1e-07;
Matches 143; Conservative 118; Mismatches 246; Indels 197; Gaps 25;
QY 307 QLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKXKASEPGKVIVLVNKV 366
Db 86 EVRDYQYTIHVHSL-FQNTLCAIPTGMGKTFIASTVMLNFRWTKA---KIIFTAPTR 140
QY 367 LLVEQLRKEFPFLKWKYRVIGLSG-----TOLKISFPEVVKSCDIIISTAQILENS 420
Db 141 PLVAQQ-----IKACLGITGPSDQTAILLDKSKRNREIWKRVFFATPQVEND 192
QY 421 LLNLENGEDAGVQLSDFLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKNKVP 480
Db 193 L-----KRGVLDPKDIVCLVIDEHRATGSSAYTNVVKF-----IDRFNSY-- 234
QY 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIK-TVKENLDQLKNQIQEPCCKF 539
Db 235 --RLALATATP-----ASDLEGVQEVVNNLDISKIEIRTEESMDIVKMKRKKKEKI 284

QY 540 AIADATREDPFKEKLLLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQM-----EKAA 590
Db 285 EV-----PLLEIEDIE-----QLGMAVKPLVQAIELGIYECDPSQINAF 327
QY 591 KGNRKERVCA-----EHLRKYNEAL-----QINDTIRMIDAYTHLETFYNEEKDKKFAV 640
Db 328 KAMQSQKIIANPTIPIGKWRNFFILQLNNVQMLKRLKIY-GIRTFNYFQNK----- 382
QY 641 IEDDSDEGDDDEYCDGDEDEDLKKPLKLDLTDRLMTLFFEN-NKMLKRLAEN-----PE 695
Db 383 -----CTEFTTKYNLKK-----STNKIAAEFYHPLKNIKNCQENYLSDPK 424
QY 696 YENE-KLTKLRNTIMEQYTRTESARGIIPTKTQSAAYALSQWI----- 738
Db 425 FVGHGKLQCVRDLEMDFFQKRGSDSRVITFELRESALEIVKFDISVADDDQIRPHIFIGQ 484
QY 739 -----TENEFKFAEVG-----VKAHLIGAGH 759
Db 485 ARAKEGFDEVKYTRKHAPKGRKKVERLHRQEQEKFLAEARTKRAANDKLSARRTGSSE 544
QY 760 SFEKPMTOEQKEVISKFRGTGKINLIATTVABEGLDIKECNIVIRYGLVTNEIAMVOA 819
Db 545 EAQISGMNQKQKEVIHNFVKGEYNVLVCTSIGEEGLDIGEVDLIICYDTTSSPIKNIQR 604
QY 820 RGR-ARADESTVYLVHSGSGVIEHETVNDFREKMYKAIHCQVQNMK-----PE- 867
Db 605 MGRTRKRDGKIVLLFSSNESYKFERAMEDYSTLQALISKQCIDYKKSRIIPEDIPEC 664
QY 868 -----EYAHKILELQMQSIMEKMKTKRNIATK 894
Db 665 HETLITINDENIEMEDVDVIRYATQCMGKVKVKKPKKAITK 708

RESULT 8

YQHH_BACSU
ID YQHH_BACSU STANDARD; PRT; 557 AA.
AC P54509;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical helicase yqhH.
GN YQHH OR BSU24580.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes.";
RL Microbiology 142:3103-3111(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Mael C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A., "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis." Nature 390:249-256(1997).

CC -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D84432; BAA12545.1; -.

DR EMBL; Z99116; CAB14389.1; -.

DR PIR; G69958; G69958.

DR Subtilist; BG11699; yqhH.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR000330; SNF2_N.

DR Pfam; PF00271; helicase_C; 1.

DR Pfam; PF00176; SNF2_N; 1.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00490; HELICC; 1.

DR KW Hypothetical protein; Helicase; ATP-binding; Complete proteome.

FT NP_BIND 83 90 ATP (BY SIMILARITY).

FT SITE 175 178 DEAD BOX.

SQ SEQUENCE 557 AA; 64672 MW; 8693636CC219081A CRC64;

Query Match 4.2%; Score 223; DB 1; Length 557;

Best Local Similarity 21.5%; Pred. No. 0.00045;

Matches 141; Conservative 98; Mismatches 198; Indels 218; Gaps 31;

QY 304 PELQLRPYQMEVAQPALEGKN--IIICLTGSGKTRVAVYIAKDHLKKKASEPGKIV 361

Db 56 PSFTPLPHQLEVAQKVVEKNGKAILADEVGLGKIVEAGLILKEYMIR----GLAKKILI 111

QY 362 LVNKKVLLVEQLFRKEFQPL-----KKWYRVIGSGDTQLKISFPPEVVKSCDIIISTAQ 415

Db 112 LV-PASLVSQWVKELQEKELIPAVEQKKS-----VWEQCDIVVSSID 153

QY 416 ILENSLLNLENGEDAGVOLS-DFSIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKE 474

Db 154 TAKRS-----PHREIVLSIPYDLVVIDEAE-----KLKNSKTK-- 186

QY 475 NKPVIPLQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQE 534

Db 187 -----NYEFVRNLVKK 197

QY 535 PCKKFAIADATREDPFKEKLEIMTRIQTYCQSPMSDFGTQPYEQWAIQMEKKAACKGN 594

Db 198 YC--LLLTAT--PIQNRIEIFNLVSL--LKP-GHLGSONHFQ-----EEFAKKKS 242

QY 595 RKERVCAEHLRKVNEALQINDTIRMIDA-----YHLETFYNEEKDKKFAVIEDDSDEG 649

Db 243 LEAH--EHLKDLVKNVIRN--RRHDTGLNWKQRHVETVPIQFSPSEQALYDEIS---- 293

QY 650 DDEYCDGDEDEDDLKPLKLDLDTDFLMTL-----FFENKQML-KRLAENPEYE 697

Db 294 -----RLKDSINKPASMFS-----IMTLQRECCSSREAVYMTLKKMLDQKEKQAPAD 341

QY 698 NEKLTKLRTIMEQYTRTESA-----RGIIFTKRQSAYALSQWITENEKEFAE 746

Db 342 EDTISVLIDRI-NQVTQNSKALQVVDLIKKIDDKVVIIFTEYRATQIYL-QWFLQONGISS 399

QY 747 VGVKAHLIGAGHSSEFKPMTQNEQKEVISKFTGKINLLIATTVAEGLDIKECNIVIR 806

Db 400 V-----PFRGGFKRGKDKWMDLFRGKIQLVIATEAGGGINLQFCNHMIN 445

QY 807 YGLVTNEIAMVQARGRA-RADESTYVLVAHSGSGVIEHETVNDFREKMMYKAHC----- 860

Db 446 YDLPNWMLRLEQRIHRLGQERDV---HIYNMATKH-TVEEHILKLLYEKIHLFEKVV 501

QY 861 -----VQNMKPEEYAHKILELQMQSIMEKKMKTKRNIKHYKNNPSLITFL 906

Db 502 GELDDILTQIVNNFEHLHDIL---YHSATEEMKIKMD-----NLTSFL 544

RESULT 9

CENE_HUMAN STANDARD; PRT; 2663 AA.

ID CENE_HUMAN Q02224;

AC Q02224;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Centromeric protein E (CENP-E protein).

GN CENPE.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93024922; PubMed=1406971;

RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.;

RT "CENP-E is a putative kinetochore motor that accumulates just before mitosis.";

RT Nature 359:536-539(1992).

RL [2]

RN CHARACTERIZATION.

RP MEDLINE=95196755; PubMed=7889940;

RX Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.;

RA "Mitotic HeLa cells contain a CENP-E-associated minus end-directed microtubule motor.";

RT EMBO J. 14:918-926(1995).

RN [3]

RP CHARACTERIZATION.

RX MEDLINE=98437347; PubMed=9763420;

RA Chan G.K.T., Schaar B.T., Yen T.J.;

RT "Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hBUBR1.";

RL J. Cell Biol. 143:49-63(1998).

RN [4]

RP FARNESYLATION.

RX MEDLINE=20459117; PubMed=10852915;

RA Ashar H.R., James L., Gray K., Carr D., Black S., Armstrong L., Bishop W.R., Kirschmeier P.;

RT "Farnesyl transferase inhibitors block the farnesylation of CENP-E and CENP-F and alter the association of CENP-E with the microtubules.";

RL J. Biol. Chem. 275:30451-30457(2000).

CC -!- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE KINETOCHORE MOTOR. ACCUMULATES JUST BEFORE MITOSIS AT THE G2 PHASE OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT AND/OR SPINDLE ELONGATION.

CC -!- SUBUNIT: INTERACTS WITH CENP-F AND BUBR1 KINASE.

CC -!- SUBCELLULAR LOCATION: ASSOCIATES WITH KINETOCHORES DURING CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE, AND IS QUANTITATIVELY DISCARDED AT THE END OF THE CELL DIVISION.

CC -!- SIMILARITY: Belongs to the kinesin-like protein family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----

DR EMBL; Z15005; CAA78727.1; --
DR PIR; S28261; S28261.
DR HSSP; P17119; 3KAR.
DR Genew; HGNC:1856; CENPE.
DR GK; Q02224; --
DR MIM; 117143; --
DR GO; GO:0005699; C:kinetochore; TAS.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0008350; F:kinetochore motor activity; TAS.
DR GO; GO:0000067; P:DNA replication and chromosome cycle; TAS.
DR GO; GO:0007079; P:mitotic chromosome movement; TAS.
DR GO; GO:0007080; P:mitotic metaphase plate congression; TAS.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;
KW Cell cycle; Centromere; Lipoprotein; Prenylation.
FT DOMAIN 1 335 KINESIN-MOTOR.
FT DOMAIN 336 2471 COILED COIL (POTENTIAL).
FT DOMAIN 2472 2663 GLOBULAR (POTENTIAL).
FT NP BIND 86 93 ATP (BY SIMILARITY).
FT LIPID 2660 2660 S-farnesyl cysteine.
SQ SEQUENCE 2663 AA; 312087 MW; CEFCL3880C8C8CB8 CRC64;

Query Match 4.1%; Score 217.5; DB 1; Length 2663;
Best Local Similarity 18.6%; Pred. No. 0.0072;
Matches 209; Conservative 185; Mismatches 421; Indels 307; Gaps 47;

QY 9 ENFRYLISCFRVRKMYIQVEPVDYLTFLPAEVEKEIQRTVATSGNMQAVELLSTLEK 68
Db 1040 EQQKIFSLIQEKNELOQMLSEVIAEKEQLTKDLKENIEMTIE---NQEELRLGDELKK 1096
QY 69 GVHGLGWTRFV-----EALRRGTSPLAARYVMNPDLTLPSPSFENAHDEYLLQLNLLQP 123
Db 1097 -----QGEIVAQEKNAHKAKEGELSRTCRLAEVEEKLKEKSQQLQEQQLLNQVEE 1149
QY 124 TLVDKLLVRDV--LDKMEHEELLTIE--DRNRIAAEN-NGNESGVRELLK-RIVQKE-- 175
Db 1150 MSEQKKINEIENLKNELKNKELTLEHMETERELELAQKLNNENYEEVKSITKERVKLQ 1209
QY 176 -----NWFSAFNLVRLQTG-----NNELVQELTGCDCSESNAEINL 212
Db 1210 KSFETERDHLRGYIREIATGLQTKKEELKTAHIHLKEHQETIDELRRS-VSEKTAQIINT 1268
QY 213 SQVDGPQV-----EQLLSTTVQPNLEKEVWGMENNSSESFADSSVVSSEDTLS 261
Db 1269 QDLEKSHTKLQEEIPVLHHEEQELL-----PNVKVSETQETMNELELLTEQSTTKDS-TT 1322
QY 262 LAEGSVSCLDESIGHNSNMGSDSGTMGSDSDENVAARASPEPELQRPYQMEVAQPALE 321
Db 1323 LARIEM-----ERLRLNEKF-----QESQEIKSITKERNLTKIKEALEVKHDQL- 1368
QY 322 GKNIICLPTGSGKTRVAVYIAK-----DHLDKKKKASEPGKVVIVLNVKLLVEQL 372
Db 1369 -----KEHIRETLAKIQESQSQSQSLNMKEKDNETTKI-----VSEMEQ- 1408
QY 373 FRKEFPFLKKYRVIGLSGDTQLKISFPE---VVKSCDIIISTAQILENSLLNLENG-- 427
Db 1409 FKPKDSALLRIEIEMLGLS--KRLQESHDEMKSVAKKODLQRLQEVLSQESDQLKENIK 1466
QY 428 EDAGVQLSDFSLLIIDCHHTNKEAVYNNIMRHYLMQKLKNRLKKNKPVIPLPQILGL 487
Db 1467 EIVAKHLETEELKVAHCCLKEQETIN-----ELRVNLSEKETE----- 1506

QY 488 TAPGVGGATKQAKA-----EEHILKLCANLDAFTIKTVKENLDQKNQIQEPCCKF 539
Db 1507 -----ISTIQKLEALNDKLNQIQEYIEKEEQNLKIQISEVQENVNELK-QFKEHRK-- 1558
QY 540 AIADATREDPFKEKLEIMTRIQTYCQ--MSPMSDFGTQPYEQWAIQWMEKKAACKGNR-- 595
Db 1559 --AKDSALQSIKSMLELTNRLQESQEEIQIMIKEEMKRVQEQALQTERDQLKENTKEI 1616
QY 596 -----KERVCAEHLRKYNEALQIN-----DTIRMID-AYT 624
Db 1617 VAKMKEQKEYQFLKMTAVNETQCKMCEIHLKEQFETQKLNLENIETENIRLTQILHE 1676
QY 625 HLETFYNEEKDKKPAVIEDDSDEGDDDEYCDGDEDDDLKKPLKD----- 670
Db 1677 NLEEMRSVTKER-----DDLRSVEETLKVVERDQKLENLRETITRDLEKQELKIVHML 1730
QY 671 ----ETDRFLMTLFFENNMKMLKRLAENPEYENEKL-----TKLRNTIMEQYTRTE 716
Db 1731 KEHQETIDKLRGIVSEKTNESINMQDKLHSDNDALKAQDLKIQEELRIAHMHLKEQOETI 1790
QY 717 ESARGIIFTKTROSAYALSQWITEN----EKFAEVGVKAHHLIGAGHS-----SEFK 764
Db 1791 DKLRGIVSEKTKLSNMQKDLNSNAKLQEKIQELKANEHQILITLKKDVNETQKVVSEME 1850
QY 765 PMTQ--NEQKEVISKFTGKINLLIATTVABEGLDIKECNIVIRYGLVTNEIAMVQARGR 822
Db 1851 QLKQIKDQSLTSLKLEIENLNL--AQELHENLEEMK-----SVMKERDN 1893
QY 823 ARADESTYVLVAHSGSGVI-----EH-ETVNDFREKMMYKAH 859
Db 1894 LRRVEETLKLRLDQLKESLQETKARDLEIQELKTARMLSKHEKETVDKLRKISEKTIQ 1953
QY 860 CVQNMK-----PEEYAHKILELQMSIM-----EKKMKTNRNIKHYKNPNSLI 903
Db 1954 ISDIQKDLKSKDELQKQIQELQKKELQLLRVKEDVNMSSHKKINEMEQLKQFEPN---- 2009
QY 904 TFLCKNCSVLACSGEDIHVIKMHVNMTPPEFKELYIVRENK 945
Db 2010 -YLCK-----CEMDNFQTLTKLHE-----SLEEIRIVAKER 2039

RESULT 10
USO1_YEAST STANDARD; PRT; 1790 AA.
AC P25386;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Intracellular protein transport protein USO1.
GN USO1 OR INT1 OR YDL058W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RX MEDLINE=91185402; PubMed=2010462;
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA Yamasaki M.;
RT "A cytoskeleton-related gene, usol, is required for intracellular
RT protein transport in Saccharomyces cerevisiae.";
RL J. Cell Biol. 113:245-260 (1991).
RN [2]
RP SEQUENCE OF 782-1790 FROM N.A.
RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA Kendrick K.E.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-8 FROM N.A.
RA Bai Y., Symington L.S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

Db 701 C-----PLS-----QQFCGKERNSLKKA 721
QY 596 KERVCAEH-LRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDDEYC 654
Db 722 MNDACYEKVLESINEGQI-----IVFVHSRKET----- 750
QY 655 DGEDEDDLLKPKLKDDETRFLMTLFFENNKMRLKLAENPEYENKLTCLRNTIMEQYTR 714
Db 751 -----SRTATWLKNKFAEN-----ITHKLT 772
QY 715 TEESARGIIFTKTRQSAVALSOWITENEKFAEYGVKKAHLLIGAGHSSEFKPMTQNEQKEV 774
Db 773 NDAGSKQILKT-----EAAVLDPSL---RKLIESGIGTHH---AG-----LTRSDRLS 816
QY 775 ISKFTGKINLLIATTVABEGLDIKECNIIVIRYGLV-----TNEIAMVQARG 821
Db 817 EDLFDAGLLQVLVCTATLAWGNLPAHTVIKGTDVYSPKGSWEQLSPQDVLMGLRAG 876
QY 822 RARADESTYVLVAHSGSGVIEHETVNDFREKMMYKAHCVQNMKPEYAHKILELQMSI 881
Db 877 RPRYD--TF-----GEGII-----ITQSNVQYLSVLNQQLPIESQFVSKLVNLAEV 924
QY 882 MEKMKTKRN 891
Db 925 VAGNIKCRND 934

RESULT 13
SBCC CLOAB STANDARD; PRT; 1163 AA.
AC Q97FKL;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Nuclease sbccD subunit C.
GN SBCC OR CAC2736.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum";
RL J. Bacteriol. 183:4823-4838(2001).
CC -!- FUNCTION: SbccD cleaves DNA hairpin structures. These structures
CC can inhibit DNA replication and are intermediates in certain DNA
CC recombination reactions. The complex acts as a 3'->5' double
CC strand exonuclease that can open hairpins. It also has a 5'
CC single-strand endonuclease activity (By similarity).
CC -!- SUBUNIT: Heterodimer of sbcc and sbccD (By similarity).
CC -!- SIMILARITY: Belongs to the SMC family. Sbcc subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE007771; AA080682.1; --
DR PIR; G97236; G97236.
DR InterPro; IPR003439; ABC_transporter.
KW Hydrolase; Nuclease; Exonuclease; Endonuclease; DNA replication;
KW DNA recombination; ATP-binding; Coiled coil; Complete proteome.

FT NP BIND 35 42 ATP (POTENTIAL).
FT DOMAIN 197 415 COILED COIL (POTENTIAL).
FT DOMAIN 446 1003 COILED COIL (POTENTIAL).
SQ SEQUENCE 1163 AA; 135507 MW; CESFOBD2215D7A92 CRC64;
Query Match 3.7%; Score 197.5; DB 1; Length 1163;
Best Local Similarity 18.8%; Pred. No. 0.022;
Matches 197; Conservative 187; Mismatches 399; Indels 263; Gaps 45;
QY 31 VLDYLTFLPAEVKEQIQTIVATSGNMQAVELLSTLEKGVHGLWTRFVEALRRTCSP 89
Db 44 ILDSITLSLYGEVARKSSNFMNTNCSLNVSEFQISGKEIKRYLVEREF---RDNKT 99
QY 90 LAARYMNPETDLPSPSFENAHDEYLLQLNLLQPTLVKLLVRDVLDKCMBEELTIEDR 149
Db 100 GSVRSKSAKIVDITGDEVEVLEEG-----AKSVNEKCEIIGLSLDDF 142
QY 150 NRIAAAE-----NGNESGVRELLKRIVQKENW-----FSAFLNVLQ-TGNNELV 194
Db 143 TRTVLPQGFSEFLKLEGER--RNMLERLNLQYGEDELSEFKLARKIRKEREKENVLV 200
QY 195 QELTG-----SDCESNAETENLSQVDGPQVEQLLSTTVQPNLEKEVWGMENSSSESF 249
Db 201 GELKGYENINEDVLKERRELLKENNDFNEASKEYLKAEEYNEGKEVWGLIEBEKVR 260
QY 250 ADSSVSESDTSLAEGSVSCLDESGLHNSNMGSDSGTMGSDSEENVAARASPEPELQLR 309
Db 261 VRKDLMEKKD-----EIDLKEKRLG-ESSSKVK 289
QY 310 PY-----QMEVAQPALEGKNIICLPTSGSKTRVAVYIAKDLHDKKKKASEPGKVI 360
Db 290 PYIDNYENTLKQIDILKEILSRE-----NTMKAISLEKEDMEKKLSIAKONKEK 339
QY 361 VL-----VNVKLLVEQLFRKEFPFLK-KWYRVIGLSDTQLKISFPEV-----VKSCDII 410
Db 340 ALPKFMKHHIILDAIKEKOLLNLIKLEKRLQGGKIEKLSLEASNKEELIKQNIKIDSL 399
QY 411 ISTAQILENSLLNLENGEDAGVQLSD--FSLIIDE-CHHTN----- 449
Db 400 TLKIQNLESKIDNLKVPPEYKKNINEGIFLLRNYDEKLKHKNKGLDCLDKFQVDFEAKS 459
QY 450 -KEAVYNNIMRHYLMQKLKNNRLKKNKPVIPLPQILGLTASPGVGA-TKQAKAEHIL 507
Db 460 KKEMLFNKLEERSKLDITYTKLQDLNKKD-FPKDDVL-LTFQEKLNDSRQKWAKYSEYNE 517
QY 508 KLCANLDAF-----TIKTVKENLDQKNQIQ-----EPCKKFAIADATRE----- 547
Db 518 SLKASLRVVENSEQVLRITKKEEMTKLEDKISKVNKIESLEFENMAHVLREKLKSGEACP 577
QY 548 -----DPFKEKLLEIMTRIQTQYCMSPMSDF-GTQPYE-----QWAIOMEKKA 590
Db 578 VCGSVHHIKBGFKEVDLKALETLK-----SELEGFKKKRFENEIEMCEASIKVEEKNI 632
QY 591 KKGK-----RKERVCAEHLRKYN-EALQINDTIR-MIDAYTHLET 629
Db 633 KKLNESINNLGEEFKEVSLESMEKKFNLYLKEKVNKFNLEKIQLDNLIKOLSESNKIEVE 692
QY 630 YNEEK-----DKKFAVIEDSDDEGDDDEYCDGEDEDDLLKPLKLDTRFLMTLFFEN 683
Db 693 YQKEKTVEKQCEKRIVDLKELEB-AIKEFNEVAYTIENLKAELKIQD-----FKFEM 744
QY 684 NKMLKRLAENPEYENKLTCLRNTIMEQYTRTESARGIIFTKTRQSAVALSOWITENEK 743
Db 745 KEILEKERVVEAAGE-INKLRLNLNIRHTEKEQ-----LMDKCSLKEELSKNK- 793
QY 744 FAEVGVKAHLLIGAGHSSEFKPMTQNEQKEVSKFRTGKINLLIATTVABEGLDIKECNI 803
Db 794 -AEL-----KEKDKIINEKIELI-KNKVGVDNLNLYELKEKIEG-TIKK--I 834
QY 804 VIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDFREKMMYKAHCVQN 863
Db 835 EEQYNLCDDKMMNEIEDKYRKCSD--IIKTHSNLSLKKRKVND-----IDKLNK 882

Db 312 SVALQ-----EDPSAEQTVCDKVRQLEDLSKELESQHSILKDEVYMNLL----- 356

QY 286 TMGSDSDSEENVAAASPEPE-----LQLRPYQMEVAQPALEGKNIICLPTGS 333

Db 357 KLEMDAQHIKDEFFHEREDLEFKINELLAKKEQGYVVEKLKYEREDLNRLQCCAVEQ 416

QY 334 GKTRVAVYIAKDLKDKKASEPGKVIIVLVNKVLLVEQLFRKEFPFLKKWYRVIGLSGD 393

Db 417 HNKEIQ-RLQEHH--QKEVSE-----LSETFISGEKEKALMFEIQGLKEQ 460

QY 394 TQ-LKISFPVVKSCDIIISTAQILENSLLNLENGEDAGVQLSDFSLSIIIDECHHTNKEA 452

Db 461 CENLQHEKEQEVVLYNESLREMMELQTEL-----GESACKISQEF-----ETMKQQ 506

QY 453 VYNNIMRHVLMQKLN-----NRLKKENKPVIPLPQILGLTASPGVGGATKQA 500

Db 507 QASDV--HELQOKLRSAFNEKDALLETVNRLQGENEKLLSQELV-----PELESTIKNL 559

QY 501 KA-----BEHILKLCANLDAF--TIKTVKENLDQL-----KNQIQ 533

Db 560 QADNSMYLASLGQKDTMLQLELAKISSLAKEKODFISKIKTSHEEMDDLHQKWEREQRLS 619

QY 534 EPCKKFAIADATREDPFKEKLEIMTRIQTQYCMSPMSD-----FGTQPYEQWAIQMEK 587

Db 620 VELREAAGQAAQHNSLRQVRSELGTGKIDELVREKSDNDQSIIVQMKTMTEDEALSSKI 679

QY 588 KAAKGNRKERVCAEHLRKVNEALQINDTIRMIDA---YTHLETFYNEEKDKKPAVIEDD 644

Db 680 KSLYEENN-----RLHSEKAQLSRDLALQAQDFAHKEHVAFEFKKLQLMV---- 726

QY 645 SDEGGDDEYCDGEDEDLKKPLKLDDETR--FLMTLFFENNMMLKR--LAENPEYENEKL 701

Db 727 -----EERDLNKLLENEQVQKSFVKTQLYEYLKQLRASILEEN---EEDV 770

QY 702 TKLRNTIMEQYTRTESARGIIFTKTROSVAVALSQWITENEKFAEVGVKAHLIGAGHSS 761

Db 771 VKLIQAVGESLVKVEEHNLFV-----EYDARVLELENK-----IKCLQEDSAVQCE 818

QY 762 EFKPMTQNEQKEVISKFRGTGKINL---LIATTVAEEG--LDIKECNIVIRYGLVTNEIAM 816

Db 819 ELRTLVRDSEQE-----KILLRKELDAVTSAKEALQLDLLEMK-----NTNEKAS 863

QY 817 VQARG-RARADESTYVLVAHSGSG-----VIEHET-----VNDER----- 850

Db 864 LENQTLSTQVEELSQTL--HSRNEVHDEKVLVIEHENLRLLLKQRESELDVRAELILK 921

QY 851 -----EKMVKAIHCVQNMKPEYAHKI 873

Db 922 DSLEKSPSVKDQLSLVKELEEKIESLEKESKDKDEKISKIKLVAVKAKKELDSNRKEAQT 981

QY 874 LELQMQSIMEKK-----MKTKRNIKHYKNPNSLITFLCKNCVSLACSGEDIHVIEKMH 927

Db 982 LREELESVRSEKDRLSASMKEFLQGAESYK---SLLLEYDKQSEQLDVEKERAHNFER-- 1036

QY 928 HV-----NMTPEFKELYIVRENKALQKKADYQINGEI---ICKCGQAWGTMVHK 975

Db 1037 HIEDLTQLRNSTCOYERL--TSDNEDLLARIETLQANAKLLEAQILEVQKAKG--VVEK 1092

QY 976 GLDLPCL-----KIRNFV-----VVFKNNSTKKQYKMWVE 1005

Db 1093 ELDAELQKEQKIKHEVSVTVNELEELQLQFQKEKQLQKTMQ 1134

RESULT 15

GOA4 HUMAN

ID GOA4_HUMAN STANDARD; PRT; 2230 AA.

AC Q13439; Q13270; Q13654; Q14436;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Golgi autoantigen, golgin subfamily A member 4 (Trans-Golgi p230) (256

DE kDa golgin) (Golgin-245) (72.1 protein).

GN GOLGA4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.; AND ALTERNATIVE SPLICING.

RX MEDLINE=96215236; PubMed=8626529;

RA Erlich R., Gleeson P.A., Campbell P., Dietzsch E., Toh B.-H.;

RT "Molecular characterization of trans-Golgi p230: a human peripheral

RT membrane protein encoded by a gene on chromosome 6p12-22 contains

RT extensive coiled-coil alpha-helical domains and a granin motif.";

RL J. Biol. Chem. 271:8328-8337(1996).

RN [2]

RP SEQUENCE FROM N.A.

RA Seelig H.P.;

RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 131-2230 FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=96125112; PubMed=8537393;

RA Fritzler M.J., Lung C.-C., Hamel J.C., Griffith K.J., Chan E.K.L.;

RT "Molecular characterization of golgin-245, a novel Golgi complex

RT protein containing a granin signature.";

RL J. Biol. Chem. 270:31262-31268(1995).

RN [4]

RP SEQUENCE OF 524-672 FROM N.A.

RC TISSUE=Gastric fundus;

RA Balague C.;

RL Thesis (1994), Instituto municipal de investigacion medica, Spain.

CC -!- FUNCTION: May play a role in vesicular transport from the trans-

CC Golgi.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic, peripheral membrane protein

CC associated with the trans-Golgi network.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=4;

CC Comment=Additional isoforms seem to exist;

CC Name=1;

CC IsoId=Q13439-1; Sequence=Displayed;

CC Name=2;

CC IsoId=Q13439-2; Sequence=VSP_004272, VSP_004273;

CC Name=3;

CC IsoId=Q13439-3; Sequence=VSP_004274;

CC Name=4;

CC IsoId=Q13439-4; Sequence=VSP_004275;

CC -!- DOMAIN: Extended rod-like protein with coiled-coil domains.

CC -!- DISEASE: Antibodies against GOLGA4 are present in sera from

CC patients with Sjogren's syndrome (SS) [MIM:270150]. Sera from

CC patients with SS often contain antibodies that react with normal

CC components of the Golgi complex.

CC -!- DISEASE: Antibodies against GOLGA4 are found in sera from

CC hepatitis B patients.

CC -!- SIMILARITY: Belongs to the golgin family.

CC -!- SIMILARITY: Contains 1 GRIP domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U41740; AAC50434.1; -.

DR EMBL; X82834; CAA58041.1; -.

DR EMBL; U31906; AAC51791.1; -.

DR EMBL; X76942; CAA54261.1; -.

DR Genew; HGNC:4427; GOLGA4.

DR MIM; 602509; -.

DR MIM; 270150; -.

DR GO; GO:0005802; C:Golgi trans face; TAS.

DR GO; GO:0016192; P:vesicle-mediated transport; TAS.

DR InterPro; IPR000237; GRIP_domain.

DR Pfam; PF01465; GRIP; 1.

DR PROSITE; PS50913; GRIP; 1.
KW Antigen; Golgi stack; Coiled coil; Alternative splicing.
FT DOMAIN 133 2185 COILED COIL (POTENTIAL).
FT DOMAIN 252 2096 GLU-RICH.
FT DOMAIN 2168 2215 GRIP.
FT VARSPLIC 2154 2185 TPYKGNLYHTDVSLEGEPTTEFEYLRKVLPEY -> HLTKV
FT AICTIRMSHLENLPNLSICEKFLSI (in isoform 2).
FT /FTid=VSP 004272.
FT Missing (in isoform 2).
FT /FTid=VSP 004273.
FT Missing (in isoform 3).
FT /FTid=VSP 004274.
FT FTSPRSGLF -> SWLRSSS (in isoform 4).
FT /FTid=VSP 004275.
FT R -> K (IN REF. 3).
FT Y -> H (IN REF. 3).
FT T -> A (IN REF. 3).
FT K -> E (IN REF. 3).
FT T -> A (IN REF. 3).
FT K -> E (IN REF. 3).
FT K -> N (IN REF. 3).
SQ SEQUENCE 2230 AA; 261139 MW; 3BB733DB1EA86134 CRC64;

Query Match
Best Local Similarity 3.7%; Score 197; DB 1; Length 2230;
Matches 185; Conservative 212; Mismatches 366; Indels 270; Gaps 46;

QY 15 ISCFRVRKMYIQVEPVDLYTLFLPAEVEKEIQRTVATSGNMQAVELLSTLEKGVWHLG 74
Db 384 IAQLRSRIKQ-----MTTQGEELREQEKESERAA--FELEKALSTAQK----- 425

QY 75 WTRFVEALRRGTSPLAARYMNPBELTDLPSFENAHDEYLLQLNLLQPTLVLDKLLVRDV 134
Db 426 -TEE-----ARRKLKAE--DEQIKTIKTSEB--ERISLQQLSRVQEVVDV 469

QY 135 LDKMEEELLTIEDNRRIAAENNGNESGVRELLKRIVQKENWFSAFNLVLRQTGNNELV 194
Db 470 MKKSBEQIAKLQKHEKELARKE-----QELTKKLQTREREFQEQMKVALEKSQSEYL 523

QY 195 QELTGSDCSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGMENNSSESPADSSV 254
Db 524 K-----ISQKEQESLA--LEELELQKAILT-----ESENKLRLDQQ 560

QY 255 VSESDTSLAEGSVCLDSIGHNSNMGSDGT-MGSDSDSEENVAARASPE---PELQLRP 310
Db 561 EAETRYTRILESSLEKSLQENKQSKDLAVHLEAEKNKHKEITVMVEKHKTELESK 620

QY 311 YQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKIVLVNKKVLL-- 368
Db 621 HQ-----QDAL-----WTEKLQVLKQYQTEMKLRKCEQEKETLLKKEIIFQ 665

QY 369 --VEQLFRKEFPFLKQWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENSLLNLEN 426
Db 666 AHIEEMNEKTLKLDVKQTELESLSSE-----LSEVLKA-----RHKLEELSVLKD 712

QY 427 GEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMQKLNKLNKKNPKVPILPQILG 486
Db 713 QTDKMKQELEAKM--DEQKNHHQQQVDSIIKEH----EVSQRTKALKDQINQLELL- 764

QY 497 LTASPGVGATQAKAEHEHILKLCANLDAFTIKTVKENLDQKNQIQEPCCKFAI-----A 542
Db 765 -----LKERDKHLKEHQAH-----VENLEADIKRSEGELOQASAKLDVFSYQ 807

QY 543 DATRED--PFKEKLLEIMTRIQTICQMSPMDSFGTQPYEQWAIQMEKKAACKGNKERVVC 600
Db 808 SATHEQTAYEEQLAQLQK-----LLDLETE-----RILLTKQVAEVEAQKKDVC 853

QY 601 AE---HLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDEYCDGD 657
Db 854 TELDAH-----KIQVDLMQOLEK-----QNSEMEQVKSLTQVYESKLE-----DGN 896

QY 658 EDEDDLKPLKLDDETDRFLMTLFFENNKMKLRLAE-----NPEYE----- 697

Db 897 KEQEQTKQILV--EKENMILQMREGQKKEIEILTQKLSAKEDSIHLNEEYETKFKNOEK 954
QY 698 -----NEKLTJLRNTIMEQYTRTEESARGIIFTKTRQSAYALSOWITENEKFAEVGVKAH 752
Db 955 KMEKVQKAKAQETLKKLLDQEAALK-----KELENTALELSQ---KBKQFNAKMLEMA 1007

QY 753 HLIAGHSSEFKPM--TQNEQKEVISKFTGKINLLIA-----TTVAEEGLDIKECNI- 803
Db 1008 QANSAGISDAVSRLETNQKEQIESLTVHRRELNDVISIWEKKLNQQAELQEIHEIQLQ 1067

QY 804 -----VIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGVIEHETVNDFRE 851
Db 1068 EKEQEVAEKQKILLFGCEKEM-----NKEITWL---KEEGVKQDITLNLQ 1113

QY 852 KMYKAIHCYQNMKPEE---YAH-KILE-----LQMQSIMEKKMKTNRNIKHYKNPS 901
Db 1114 QLKQKSAH-VNSLAQDETTLKAHLEKLEVDLNLKSLKENTFLOQLVELKMLAEEDKRVUS 1172

QY 902 LITFLCKNCVSLACSGEDIHVIEKMHVNMVMTPEFKELYIVRENKALQKKCADIQ----- 955
Db 1173 ELT-----SKLKTDEEFQSLKSSHE-----KSNKSLEDKSLEFKKLSEEL 1213

QY 956 -INGEIIICKGQA 967
Db 1214 AIQLDICKCKKTEA 1226

Search completed: June 2, 2004, 19:10:47
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:06:53 ; Search time 49 Seconds
(without alignments)
6600.126 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGNGYSTDENFRYLISCFA.....LPITFPNLDYSECCLFSDDED 1025

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5311	100.0	1025	4	Q9H3G6	Q9h3g6 homo sapien
2	5285	99.5	1025	4	Q9BYX4	Q9byx4 homo sapien
3	4199	79.1	1025	11	Q8R5F7	Q8r5f7 mus musculu
4	4194	79.0	1025	11	Q8K5C7	Q8k5c7 mus musculu
5	2673	50.3	693	11	Q8BZ01	Q8bz01 mus musculu
6	2633	49.6	683	11	Q8BYC9	Q8byc9 mus musculu
7	2456	46.2	468	4	Q96MX8	Q96mx8 homo sapien
8	2271.5	42.8	514	11	Q99KS4	Q99ks4 mus musculu
9	2059.5	38.8	467	11	Q8R144	Q8r144 mus musculu
10	1356	25.5	678	11	Q99J87	Q99j87 mus musculu
11	1352	25.5	678	4	Q96C10	Q96c10 homo sapien
12	1344	25.3	678	4	Q9HAM6	Q9ham6 homo sapien
13	1122	21.1	244	11	Q9D2Z5	Q9d2z5 mus musculu
14	1086.5	20.5	925	4	Q95786	Q95786 homo sapien
15	1071	20.2	221	4	Q86X56	Q86x56 homo sapien
16	1018	19.2	940	6	Q9GLV6	Q9glv6 sus scrofa

17	687	12.9	1037	5	O44165	O44165 caenorhabdi
18	678	12.8	143	11	Q8VE79	Q8ve79 mus musculu
19	555.5	10.5	472	11	Q8C7T2	Q8c7t2 mus musculu
20	555	10.5	1164	5	Q93413	Q93413 caenorhabdi
21	516	9.7	398	4	Q9NT04	Q9nt04 homo sapien
22	516	9.7	620	5	Q17545	Q17545 caenorhabdi
23	453	8.5	752	17	Q9V1Z5	Q9v1z5 pyrococcus
24	441.5	8.3	650	17	O59524	O59524 pyrococcus
25	431.5	8.1	1883	10	Q8LMR2	Q8lmr2 oryza sativ
26	417	7.9	741	17	O28814	O28814 archaeglob
27	395.5	7.4	764	17	Q8TZH8	Q8tzh8 pyrococcus
28	391	7.4	410	11	Q8C5I3	Q8c5i3 mus musculu
29	389	7.3	821	17	Q8TUH1	Q8tuh1 methanosarc
30	384	7.2	1604	10	Q7XQ14	Q7xql4 oryza sativ
31	368	6.9	864	17	Q8PX35	Q8px35 methanosarc
32	365.5	6.9	837	5	Q86L44	Q86l44 dictyosteli
33	350	6.6	513	17	Q978A0	Q978a0 thermoplasma
34	343	6.5	182	11	Q9D1X4	Q9dlx4 mus musculu
35	343	6.5	2042	10	Q9M9P8	Q9m9p8 arabidopsis
36	342	6.4	738	17	O27466	O27466 methanobact
37	332.5	6.3	1383	5	Q9VDA0	Q9vda0 drosophila
38	332	6.3	508	17	Q9HI46	Q9hi46 thermoplasma
39	313.5	5.9	784	17	Q9HWW5	Q9hmw5 halobacteri
40	309.5	5.8	1883	5	Q95ZG7	Q95zg7 dictyosteli
41	302.5	5.7	741	17	Q8TUS6	Q8tus6 methanopyru
42	302.5	5.7	1789	5	Q8T145	Q8t145 dictyosteli
43	295.5	5.6	783	3	Q9HE09	Q9he09 schizosacch
44	295	5.6	1458	10	Q8W367	Q8w367 oryza sativ
45	295	5.6	1458	10	Q7XD96	Q7xd96 oryza sativ

ALIGNMENTS

RESULT 1
Q9H3G6
ID Q9H3G6 PRELIMINARY; PRT; 1025 AA.
AC Q9H3G6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Melanoma differentiation associated protein-5.
GN MDA5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Kang D.-C., Fisher P.B.;
RT "Melanoma differentiation associated gene-5 (mda-5), an interferon
inducible gene of limited homology to RNA helicase.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF095844; AAG34368.1;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR SMART; SM00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 1025 AA; 116687 MW; 48BE75491D863741 CRC64;

Query Match 100.0%; Score 5311; DB 4; Length 1025;
Best Local Similarity 100.0%; Pred. No. 2.7e-288;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGNGYSTDENFRYLISCFA.....LPITFPNLDYSECCLFSDDED 60

Db 1 MSNGYSTDENFRYLISCFRVRVYQVEPVLDYLTFLPAEVKEQIQTIVATSGNMQAVE 60
QY 61 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPETDLPSPSFENAHDEYLQNL 120
Db 61 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPETDLPSPSFENAHDEYLQNL 120
QY 121 LQPTLVKLLVRDVLKCMEEELLTIEDNRNIAAENNGESGVRELLKRIVQKENWFS 180
Db 121 LQPTLVKLLVRDVLKCMEEELLTIEDNRNIAAENNGESGVRELLKRIVQKENWFS 180
QY 181 FLNVLRQTGNNELVQELTGSDCSSENAEIEENLSQVDPGPQVEEQLLSTTVQPNLEKEVWGM 240
Db 181 FLNVLRQTGNNELVQELTGSDCSSENAEIEENLSQVDPGPQVEEQLLSTTVQPNLEKEVWGM 240
QY 241 ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
Db 241 ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
QY 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360
Db 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360
QY 361 VLNVKVLVLEQVLFKFEFQPFLLKRWYRIGLSDGTQLKISFPEVVKSCDIIISTAQILENS 420
Db 361 VLNVKVLVLEQVLFKFEFQPFLLKRWYRIGLSDGTQLKISFPEVVKSCDIIISTAQILENS 420
QY 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480
Db 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480
QY 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
Db 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
QY 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAAGKGNRERV 600
Db 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAAGKGNRERV 600
QY 601 AEHLRYNEALQINDTIRMIDAYTHLETIFYNEEKDKFAVIEDSDDEGDDCYDGEDE 660
Db 601 AEHLRYNEALQINDTIRMIDAYTHLETIFYNEEKDKFAVIEDSDDEGDDCYDGEDE 660
QY 661 DDLKPLKLDDETDRFLMTLFFENKMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
Db 661 DDLKPLKLDDETDRFLMTLFFENKMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
QY 721 GIIFTKRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFT 780
Db 721 GIIFTKRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFT 780
QY 781 GKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARESTYVLVAHSGSGV 840
Db 781 GKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARESTYVLVAHSGSGV 840
QY 841 IEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMQSIMCKMKTNRNIAKHYNP 900
Db 841 IEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMQSIMCKMKTNRNIAKHYNP 900
QY 901 SLITFLCKNCSVLACSGEDIHVEKMHVNMTPFKELYIVRENKALQKCADYQINGEI 960
Db 901 SLITFLCKNCSVLACSGEDIHVEKMHVNMTPFKELYIVRENKALQKCADYQINGEI 960
QY 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVFKNNSTKKQYKWKVELPITFPNLDYSECCL 1020
Db 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVFKNNSTKKQYKWKVELPITFPNLDYSECCL 1020
QY 1021 FSDED 1025
Db 1021 FSDED 1025

Q9BYX4
ID Q9BYX4 PRELIMINARY; PRT; 1025 AA.
AC Q9BYX4;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE RNA helicase-DEAD box protein RH116.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Cocude C., Kolesnitchenko V., Billaut-Mulot O., Truong M.-J.,
RA Capron A., Bahr G.M.;
RT "Identification of a new RNA helicase (RH116) regulated by the
RT immunomodulator Murabutide."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY017378; AAG54076.1; -
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 1025 AA; 116671 MW; 6B4F3D14E88A7D6D CRC64;

Query Match 99.5%; Score 5285; DB 4; Length 1025;
Best Local Similarity 99.5%; Pred. No. 7.7e-287;
Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSNGYSTDENFRYLISCFRVRVYQVEPVLDYLTFLPAEVKEQIQTIVATSGNMQAVE 60
Db 1 MSNGYSTDENFRYLISCFRVRVYQVEPVLDYLTFLPAEVKEQIQTIVATSGNMQAVE 60
QY 61 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPETDLPSPSFENAHDEYLQNL 120
Db 61 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPETDLPSPSFENAHDEYLQNL 120
QY 121 LQPTLVKLLVRDVLKCMEEELLTIEDNRNIAAENNGESGVRELLKRIVQKENWFS 180
Db 121 LQPTLVKLLVRDVLKCMEEELLTIEDNRNIAAENNGESGVRELLKRIVQKENWFS 180
QY 181 FLNVLRQTGNNELVQELTGSDCSSENAEIEENLSQVDPGPQVEEQLLSTTVQPNLEKEVWGM 240
Db 181 FLNVLRQTGNNELVQELTGSDCSSENAEIEENLSQVDPGPQVEEQLLSTTVQPNLEKEVWGM 240
QY 241 ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
Db 241 ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
QY 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360
Db 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360
QY 361 VLNVKVLVLEQVLFKFEFQPFLLKRWYRIGLSDGTQLKISFPEVVKSCDIIISTAQILENS 420
Db 361 VLNVKVLVLEQVLFKFEFQPFLLKRWYRIGLSDGTQLKISFPEVVKSCDIIISTAQILENS 420
QY 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480
Db 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480
QY 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
Db 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
QY 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAAGKGNRERV 600

Db	541	IADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAKEGNRKESVC	600
QY	601	AHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKOKKFAVIEDDSDEGGDDDEYCDGDEDE	660
Db	601	AHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKOKKFAVIEDDSDEGGDDDEYCDGDEDE	660
QY	661	DDLKPLKLDLDETDRFLMTLFFENNMKRLKLAENPEYENKLTKLNTIMEQYTRTEESAR	720
Db	661	DDLKPLKLDLDETDRFLMTLFFENNMKRLKLAENPEYENKLTKLNTIMEQYTRTEESAR	720
QY	721	GIIFTKTQSAAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVISKFT	780
Db	721	GIIFTKTQSAAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVISKFT	780
QY	781	GKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSV	840
Db	781	GKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSV	840
QY	841	IEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMSIMEKMKTKRNIKHYKNNP	900
Db	841	IERETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMSIMEKMKTKRNIKHYKNNP	900
QY	901	SLITFLCKNCVSLACSGEDIHVIEKMHVNMTPFEKELYIVRENKALQKCCADYQINGEI	960
Db	901	SLITFLCKNCVSLACSGEDIHVIEKMHVNMTPFEKELYIVRENKALQKCCADYQINGEI	960
QY	961	ICKGQAWGTMVHKGLDLPCLKIRNFVVFVFNKNSSTKKQYKKWVELPITFPNLDYSECCL	1020
Db	961	ICKGQAWGTMVHKGLDLPCLKIRNFVVFVFNKNSSTKKQYKKWVELPITFPNLDYSECCL	1020
QY	1021	FSDED 1025	
Db	1021	FSDED 1025	

RESULT 3

Q8R5F7	PRELIMINARY;	PRT; 1025 AA.
ID	Q8R5F7	
AC	Q8R5F7	
DT	01-JUN-2002 (TrEMBLrel. 21, Created)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	HELICARD.	
GN	MDA5 OR 9130009C22RIK.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Kovacsovichs M., Hofmann K., Tschopp J.;	
RT	"HELICARD, a novel CARD-containing helicase.";	
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AY075132; AAL77205.1; -.	
DR	MGI; MGI:1918836; Mda5.	
DR	GO; GO:0005524; F:ATP binding; IEA.	
DR	GO; GO:0008026; F:ATP dependent helicase activity; IEA.	
DR	GO; GO:0016787; F:hydrolase activity; IEA.	
DR	GO; GO:0003676; F:nucleic acid binding; IEA.	
DR	InterPro; IPR001410; DEAD.	
DR	InterPro; IPR001650; Helicase_C.	
DR	Pfam; PF00270; DEAD; 1.	
DR	Pfam; PF00271; helicase C; 1.	
DR	SMART; SM00487; DEXDC; 1.	
DR	SMART; SM00490; HELICC; 1.	
KW	ATP-binding; Helicase; Hydrolase.	
SQ	SEQUENCE 1025 AA; 115971 MW; 708FCAC690C9F6D8 CRC64;	

Query Match 79.1%; Score 4199; DB 11; Length 1025;
Best Local Similarity 79.7%; Pred. No. 3.6e-226;
Matches 818; Conservative 75; Mismatches 131; Indels 2; Gaps 2;

QY	1	MSGYSTDENFRYLISCFRFRVQMYIQVEPVLDTLFLPAEVKEIQRTVATSGNMQAVE	60
Db	1	MSIVCSAEDSFRNLILFFRPLKMYIQVEPVLDTLFLPAEVKEIQRTVATSGNMQAVE	60
QY	61	LLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPBLTDLSPSFENAHDEYQLNL	120
Db	61	LLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPBLTDLSPSFENAHDEYQLNL	120
QY	121	LQPTLVLDKLLVRDLVKCMEEHLLTIEDNRRIAAENNGNESGVRELLKRIVQKENWFS	180
Db	121	LQPTLVLDKLLVRDLVKCMEEHLLTIEDNRRIAAENNGNESGVRELLKRIVQKENWFS	180
QY	181	FLNVLRTGNNELVQELTSGDCSESAEIEENLSQVDPQVEEQLLSTTVQPNLEKEVWGM	240
Db	181	FLNVLRTGNNELVQELTSGDCSESAEIEENLSQVDPQVEEQLLSTTVQPNLEKEVWGM	240
QY	241	ENNSSESSFADSSVSESDTSLAEGSVSCLDESIGHNSNMGSDSGTGMGSDSEENV	300
Db	241	ENNSSESSFADSSVSESDTSLAEGSVSCLDESIGHNSNMGSDSGTGMGSDSEENV	300
QY	300	ASPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKV	359
Db	301	VSPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKV	360
QY	360	IVLVNKKVLLVEQLFRKEFQFPLKKWYRVIIGSGDTQLKISFPEVVKSCDIISTAQILEN	419
Db	361	IVLVNKKVLLVEQLFRKEFQFPLKKWYRVIIGSGDTQLKISFPEVVKSCDIISTAQILEN	420
QY	420	SLINLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMOKLNNRLKKNKPKVI	479
Db	421	SLINLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMOKLNNRLKKNKPKVI	480
QY	480	PLPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKF	539
Db	481	PLPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKF	540
QY	540	AIADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAAGNRRKRV	599
Db	541	VIADDTRENPFKEKLEIMASIQTQYQMSPMDSFGTQPYEQWAIQMEKKAAGNRRKRV	600
QY	600	CAEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKOKKFAVIEDDSDEGGDDDEYCDGDE	659
Db	601	CAEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKOKKFAVIEDDSDEGGDDDEYCDGDE	659
QY	660	EDLKKPLKLDLDETDRFLMTLFFENNMKRLKLAENPEYENKLTKLNTIMEQYTRTEESA	719
Db	660	EDLKKPLKLDLDETDRFLMTLFFENNMKRLKLAENPEYENKLTKLNTIMEQYTRTEESA	719
QY	720	RGIIFTKTQSAAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVISKFR	779
Db	720	RGIIFTKTQSAAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVISKFR	779
QY	780	TGKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSG	839
Db	780	TGKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSG	839
QY	840	VIEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMSIMEKMKTKRNIKHYKNN	899
Db	840	VIEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMSIMEKMKTKRNIKHYKNN	899
QY	900	PSLITFLCKNCVSLACSGEDIHVIEKMHVNMTPFEKELYIVRENKALQKCCADYQINGE	959
Db	900	PSLITFLCKNCVSLACSGEDIHVIEKMHVNMTPFEKELYIVRENKALQKCCADYQINGE	959
QY	960	ICKGQAWGTMVHKGLDLPCLKIRNFVVFVFNKNSSTKKQYKKWVELPITFPNLDYSECCL	1019
Db	960	ICKGQAWGTMVHKGLDLPCLKIRNFVVFVFNKNSSTKKQYKKWVELPITFPNLDYSECCL	1019
QY	1020	FSDED 1025	
Db	1020	FSDED 1025	

```
RESULT 4
Q8K5C7
ID Q8K5C7 PRELIMINARY; PRT; 1025 AA.
AC Q8K5C7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Melanoma differentiation associated gene 5-like protein.
GN MDA5 OR 9130009C22RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Kang D.-C., Fisher P.B.;
RT "Murine homolog of melanoma differentiation associated gene-5 - a DEXH
RT group RNA dependent ATPase.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF374384; AAM21359.1; -.
DR MGD; MGI:1918836; Mda5.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICc; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 1025 AA; 115941 MW; 8338A8510C5ED3A8 CRC64;

Query Match 79.0%; Score 4194; DB 11; Length 1025;
Best Local Similarity 79.6%; Pred. No. 6.9e-226;
Matches 817; Conservative 75; Mismatches 132; Indels 2; Gaps 2;

QY 1 MSGYSTDENFRYLISCFRARKVMYIQVEPVDLYTLFLPAEVKEIQRTVATSGNMQAVE 60
Db 1 MSIVCSAEDSFRNLILFFRPRLKMYIQVEPVDLHLIFLSAETKEQLKINTCGNTSAE 60

QY 61 LLLSTLEKGVWHLGWTREFVEALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL 120
Db 61 LLLSTLEQGWPLGTQMFVEALEHSGNPLAARYVKPTLTLDLPSPSETAHDECLHLLTL 120

QY 121 LQPTLVDKLLVRDVLDKCMEEBELLTIEDNRPIAAENNGNESGVRELLKRIVQKENWPSA 180
Db 121 LQPTLVDKLLINDVLTDFEKGLLTVEDNRPIISAAGNSGNESGVRELLRRIVQKENWFS 180

QY 181 FLNVLRQTGNELVQELTGSDCSNAELENLSQVDGPQVEQLLSTTVQPNLEKEVWGM 240
Db 181 FLVLRQTGNLALFQELTGGCPEDNTDLANSRRDGPAADECLPAVDSESSLETRANNV 240

QY 241 ENNSSESPADSSVVSSEDTSLAEGSVSLDESGLGHNSNMGSDSGTMGSDSDEENV-AAR 299
Db 241 DDILPEASCTDSSVTTESDTSLAEGSVSCFDESGLGHNSNMGRDSTMGSDSDSVIOTKR 300

QY 300 ASPEPELQRPYQMEVAQPALEGKNIITCLPTGSGKTRVAVYIAKDHLDKXKASEPGKV 359
Db 301 VSPEPELQRPYQMEVAQPALDGKNIITCLPTGSGKTRVAVYITKDHLDKXKASESGKV 360

QY 360 IVLVNKVLLVEQLFRKEFPFLKKWYRVIGLSDGTQLKISFPEVVKSCDIIISTAIQLEN 419
Db 361 IVLVNKVMLAEQLFRKEFPYLPYKKWYRIIGLSDGTQLKISFPEVVKSYDVIISTAIQLEN 420

QY 420 SLINLENGEDAGVQLSDFSLIIIDECHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPKVI 479
Db 421 SLINLESGDDGQVLSDFSLIIIDECHTNKEAVYNNIMRYLKQKLNNRLKKNKPKPAI 480

QY 480 PLPQILGLTASPGVGATKQAKAEHLKLCANLDAFTIKTVKENLDQLKNQIQEPCCKF 539
```

```
Db 481 PLPQILGLTASPGVGAACKQSEAEKHILNICANLDAFTIKTVKENLQLKHIKEPCCKF 540
QY 540 AIADATREDPPKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAACKGNRKERV 599
Db 541 VIADDTRENPEKLEIMASIQTYCQKSPMSDFGTQHYEQWAIQMEKKAACKGNRKDRV 600
QY 600 CAEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDDEGGDDEYCDGED 659
Db 601 CAEHLRKYNEALQINDTIRMIDAYSHLEAFYDEKEKKFAVL-NDSDKSDDEASSCNDQL 659
QY 660 EDDLKPLKLDLTDRLMTLFFENNKMRLKLAENPEYENKLTKLNTIMEQYTRTEESA 719
Db 660 KGDVKKSLKLDLTDRLMTLFFENNKMRLKLAENPEYENKLTKLNTIMEQYTRSESS 719
QY 720 RGIIFTKTROSAYALSQWITENEKEFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKER 779
Db 720 RGIIFTKTROSAYALSQWIMENAKFAEVGVKAHHLIGAGHSSEFKPMTQTEQKEVISKER 779
QY 780 TGINLLIATVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYLVVHSGSG 839
Db 780 TGEINLLIATVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYLVVTSVSGSG 839
QY 840 VIEHETVNDREKMMYKAHCVQNMKPEYAHKILELQMSIMEKMKTKRNIKHYKNN 899
Db 840 VTEREIVNDREKMMYKAINRVQNMKPEYAHKILELQVQSILEKMKVKRSIAKQYNDN 899
QY 900 PSLITFLCKNCVSLACSGEDIHVEIKMHVNMTPFEKELYIVRENKALOKKCADYQINGE 959
Db 900 PSLITLLCKNCMLVCSGNIHVEIKMHVNMTPFEKGLYIVRENKALOKKFADYQINGE 959
QY 960 IICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSKKQYKKNWVELPITFPNLDYSECC 1019
Db 960 IICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSPKQYKKNWVELPIRFPDLDYSEYC 1019
QY 1020 LFSDED 1025
Db 1020 LYSDED 1025

RESULT 5
Q8BZ01
ID Q8BZ01 PRELIMINARY; PRT; 693 AA.
AC Q8BZ01;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to melanoma differentiation associated protein-5
DE (Fragment).
DE MDA5 OR 9130009C22RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Vagina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK037057; BAC29687.1; -.
DR MGD; MGI:1918836; Mda5.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
DR InterPro; IPR001410; DEAD.
DR Pfam; PF00270; DEAD; 1.
DR SMART; SM00487; DEXDC; 1.
FT NON TER 693
SQ SEQUENCE 693 AA; 77757 MW; BB11B458B5EE019D CRC64;

Query Match 50.3%; Score 2673; DB 11; Length 693;
```


OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025508; AAK25508.1; --
DR GO:0005524; F:ATP binding; IEA.
DR GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO:0016787; F:hydrolyase activity; IEA.
DR GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00490; HELIC_C; 1.
KW Hypothetical protein; ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 467 AA; 53966 MW; F3B0D976778F0442 CRC64;

Query Match 38.8%; Score 2059.5; DB 11; Length 467;
Best Local Similarity 84.6%; Pred. No. 4.2e-107;
Matches 396; Conservative 28; Mismatches 43; Indels 1; Gaps 1;

QY 558 MTRIQTQYCMSPMSDFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQINDTI 617
Db 1 MASIQTQYCMSPMSDFGTQHYEQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQINDTI 60
QY 618 RMIDAYTHLETFYNEEKDKKFAVIEDSDDEGDDDEYCDGDEDDDLKKPLKLDDETRFLM 677
Db 61 RMIDASHLETFYDEKEKFAVL-NDSDSDDEASSCNDQLKGDVKKSLKLDDETRFLM 119
QY 678 TLFFENNKKLAENPEYENKLTNRTIMEQYTRTESARGIIFTKTRQSAVALSQW 737
Db 120 NLFFDNKKMLKLAENPKYENKLIKLNRTILEQFTRSESSRGIIFTKTRQSTVALSQW 179
QY 738 ITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFTGKINLLIATTVAEEGLD 797
Db 180 IMENAKFAEVGVKAHHLIGAGHSSEVKPMTQTEQKEVISKFTGTEINLLIATTVAEEGLD 239
QY 798 IKECNIVIRYGLVTNEIAMVQARGARADESTYLVVAHSGSGVIEHETVNDFREKMYKA 857
Db 240 IKECNIVIRYGLVTNEIAMVQARGARADESTYLVTSSSGVTEREIVNDFREKMYKA 299
QY 858 IHCVQNMKEPEYAHKILELOMQSIMKGMKTKRNIKHYKNPSLITFLCKNCSVLACSG 917
Db 300 INRVQNMKEPEYAHKILELQVQSILEKMKVKRSIAKQVNDNPSLITLLCKNCSMLVCSG 359
QY 918 EDIHVIEKMHVNMTPEFKELYIVRENKALQKKADYQINGEIIICKGQAWGTMVHKGL 977
Db 360 ENIHVIEKMHVNMTPEFKGLYIVRENKALQKKFADYQINGEIIICKGQAWGTMVHKGL 419
QY 978 DLPCLKIRNFVVFKNSTKKQYKKWVVELPITFPNLDYSECCLFSD 1025
Db 420 DLPCLKIRNFVVFKNNSPKQYKKWVVELPIRFPDLDYSEYCLYSD 467

RESULT 10
Q99J87 PRELIMINARY; PRT; 678 AA.

AC Q99J87;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (DNA segment, Chr 11, Lothar Hennighausen 2, expressed).
DE D11LGP2E OR D11LGP2.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11161808;

RA Miyoshi K., Cui Y., Riedlinger G., Lehoczy J., Zon L., Oka T.,
RA Dewar K., Hennighausen L.;
RT "Structure of the Mouse Stat 3/5 Locus: Evolution from Drosophila to
RT Zebrafish to Mouse.";
RL Genomics 71:150-155(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF317000; AAK15475.1; --
DR EMBL; AF316999; AAK15474.1; --
DR EMBL; BC029209; AAK29209.1; --
DR MGD; MGI:1931560; D11LGP2e.
DR GO:0005524; F:ATP binding; IEA.
DR GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO:0016787; F:hydrolyase activity; IEA.
DR GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
KW Hypothetical protein; ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 678 AA; 76726 MW; DC42B75A3AD376A8 CRC64;

Query Match 25.5%; Score 1356; DB 11; Length 678;
Best Local Similarity 41.7%; Pred. No. 1.4e-67;
Matches 299; Conservative 129; Mismatches 237; Indels 52; Gaps 12;

QY 306 LQLRPYQMEVAQPALEGKNIICLTGSGKTRVAVYIAKHLDKKKASEPGKVIIVLNK 365
Db 1 MELRPYQWEVILPALEGKNIICLTGAGKTRAAAFVAKRHLE---TVDRGKVVVNVNR 56
QY 366 VLLVEQLFRKEFQPFELKKWYRIGLSDTQLKISFPEVVKSCDIIISTAQILENSLLNLE 425
Db 57 VHLVSQ-HAEFRRLDKHWTVTLSGDMGSRAGFGLMARSHDLLICTAELLQALNSSE 115
QY 426 NGEDAGVQLSDFSLIIDECHTNKEAVYNNIMRHYLMQKLNNRLKKNKPVILPQIL 485
Db 116 --EDEHVELREFSLIVDECHTHKDTVNTILSRYLEQKLK---KAE-----PLPQVL 164
QY 486 GLTASPGVGATKQAKAEHEHLKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFAIADAT 545
Db 165 GLTASPGTGATKLGAIIDHILQLCANLDTCHIMSPKNCYSQLLMHNPKCKQVLDLQRR 224
QY 546 REDPFKEKLEIMTRIQTQYCMSPM-SDFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHL 604
Db 225 AQDPFGDLIKKLMNQIHQQLMEPDLKQOFGTQYEQVQVQLCKDAEAGLQEQRYALHL 284
QY 605 RKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDDEGDDDEYCDGDEDDDLK 664
Db 285 RRYNDALFIHDTVARDALDMLQDFYDRERTTKTQMVRAES-----PLPQVLAES 325
QY 665 KPLKLDDETRFLMTLFFENNKKLAENPEYENKLTNRTIMEQYTRTESARGIIF 724
Db 326 -----WLLKLFDDHKNVGLQLAARGP-ENPKLEMLERILLKQF-GSPGHTRGIIIF 373
QY 725 TKTRQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFTGKIN 784
Db 374 TRTRQTASSLLWLRQQLQPCQCTVGIKPKQMLIGAGNTSQSTHMTQDQEQVEIQEFRDGILS 433
QY 785 LLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYLVVAHSGSGVIEHE 844
Db 434 LLVATSAEEGLDIAQCNVVVRVYGLLTNEISMVQARGARAGQSVYSFLATEGSREMKRE 493
QY 845 TVNDFREKMYKAHCVQNMKEPEYAHKILELOMQSIMKGMKTKRNIKHYKNPSLIT 904
Db 494 LTNEALEVLMKAVAAVQKMDPEFKAKIRDLQOASLVKRAARAAREIQOQGFLEHVQ 553
QY 905 FLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENK-ALQKKADYQINGEIIIC- 962
Db 554 LLCINCMVAVGYGSDLRKVEGTHHVNPNFVSYYTTSQNPVIVNKVFKDWRPGGTIRCS 613
QY 963 KCGQAWGTMVHKGLDLPCLKIRNFVVFKNSTKKQYKKWVVELPITFPNLDYSECC 1019

Db 614 NCGEVWGFQMIYKSVTLPLVLKIGS--MLLETGPRGKIQAQKWSRVFSPFDPILQDC 668

RESULT 11

Q96C10 PRELIMINARY; PRT; 678 AA.

AC Q96C10;

DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Hypothetical protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Colon;

RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC014949; AAH14949.1; -

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.

DR GO; GO:0016787; F:hydrolyase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001650; Helicase_C.

DR Pfam; PF00270; DEAD; 1.

DR Pfam; PF00271; helicase_C; 1.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00490; HELIC; 1.

KW Hypothetical protein; ATP-binding; Helicase; Hydrolase.

SQ SEQUENCE 678 AA; 76612 MW; 859E1749C7313D06 CRC64;

Query Match 25.5%; Score 1352; DB 4; Length 678;

Best Local Similarity 40.7%; Pred. No. 2.3e-67;

Matches 292; Conservative 136; Mismatches 237; Indels 52; Gaps 12;

Qy 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKVIVLVNK 365

Db 1 MELRSYQWEVIMPALEGKNIILWPTGAGKTRAAAYVAKRHLE---TVDGAKVWVLVNR 56

Qy 366 VLLVEQLFRKEFQFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAIQLENSLNL 425

Db 57 VHLVTQ-HGEFFRRMLDGRWTVTTLSDGMGPRAGFHLARCHDILLICTAELLQMALTSPE 115

Qy 426 NGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKNKPVIPLPQIL 485

Db 116 --EEHVELTVFSLIVVDECHHTKDTVYNNVMSQYELKQLRAQ-----PLPQVL 164

Qy 486 GLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFAIADAT 545

Db 165 GLTASPGTGASKLDGAINHVLQLCANLDTWCIMSPQNCPPQOEHSQQPCQYNLCHRR 224

Qy 546 REDPFKEKLLLEIMTRIQTVCQSPMS-DFGTQPYEQWAIQMEKKAACKGNRKERVCAEHL 604

Db 225 SQDPFGDLLKLMDOIHDLMPLELSRKFGTQMTYEQVQVVKLSEAAALAGLQEQRVYALHL 284

Qy 605 RKNZALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDDLK 664

Db 285 RRYNDALLIHDITRAVDALAAALQDFYHREHVTKTQIL-----C----- 322

Qy 665 KPLKLDETDRFLMTLFFENNNKMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIF 724

Db 323 -----AERRLLALFDDRKNELAHLATHGP-ENPKLEMLEKILQRFQFS-SSNSPRGIIF 373

Qy 725 TKTRQSAVALSQWITENEFKFAEVGVKAAHLLIGAGHSSEFKPMTONEQKEVSKFRTGKIN 784

Db 374 TRTRQSAHSLLLWLQOQQGLQTVDIRAQLLIGAGNSQSQTHMTQDQEQVQKFDGTLN 433

Qy 785 LLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQAPGRARADESTYVLVAHSGSGVIEHE 844

Db 434 LLVATSVAEGLDIPHCVVVRVYGLLTNEISMVQARGRARADQSVYAFVATEGSRELKRE 493

Qy 845 TVNDFREKMMYKAHICVQNMKPEEYAHKILELQMSIMEKMKTKRNIATKHYKNPSLIT 904

Db 494 LINEALETLMEQAAVQKMDQAEYQAKIRDLQQAALTAKRAAQAQRENQOQFPVEHVQ 553

Qy 905 FLCKNCVSLACSGEDIHVIEKMHVNMTPFEKELY-IVRENKALQKCKADYQINGEIIICK 963

Db 554 LLCINCMVAVGHSGLRKVEGTHHVNPNFNSYNNVSRDPVWINKVFKDWKPGGVISCR 613

Qy 964 -CGQAWGTMVHKGLDLPCLKIRNFVVFVFNKNNSTKKQYKKWVLPITFNLIDYSECC 1019

Db 614 NCGEVWGLQMIYKSVKLPVLKVR--MLLETGPRGKIQAQKWSRVFSPFDFLQHC 668

RESULT 12

Q9HAM6 PRELIMINARY; PRT; 678 AA.

AC Q9HAM6;

DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Hypothetical protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RA Isogai T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;

RT "NEDO human cDNA sequencing project."

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK021416; BAB13818.1; -

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.

DR GO; GO:0016787; F:hydrolyase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001650; Helicase_C.

DR Pfam; PF00270; DEAD; 1.

DR Pfam; PF00271; helicase_C; 1.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00490; HELIC; 1.

KW Hypothetical protein; ATP-binding; Helicase; Hydrolase.

SQ SEQUENCE 678 AA; 76642 MW; 0628A21A71856EE5 CRC64;

Query Match 25.3%; Score 1344; DB 4; Length 678;

Best Local Similarity 40.6%; Pred. No. 6.6e-67;

Matches 291; Conservative 136; Mismatches 238; Indels 52; Gaps 12;

Qy 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKVIVLVNK 365

Db 1 MELRSYQWEVIMPALEGKNIILWPTGAGKTRAAAYVAKRHLE---TVDGAKVWVLVNR 56

Qy 366 VLLVEQLFRKEFQFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAIQLENSLNL 425

Db 57 VHLVTQ-HGEFFRRMLDGRWTVTTLSDGMGPRAGFHLARCHDILLICTAELLQMALTSPE 115

Qy 426 NGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKNKPVIPLPQIL 485

Db 116 --EEHVELTVFSLIVVDECHHTKDTVYNNVMSQYELKQLRAQ-----PLPQVL 164

Qy 486 GLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFAIADAT 545

Db 165 GLTASPGTGASKLDGAINHVLQLCANLDTWCIMSPQNCPPQOEHSQQPCQYNLCHRR 224

Qy 546 REDPFKEKLLLEIMTRIQTVCQSPMS-DFGTQPYEQWAIQMEKKAACKGNRKERVCAEHL 604

Db 225 SQDPFGDLLKMLMDQIHDHLEMPELSRKFGTQMVEQQVVKLSEAAALAGLQEQRVYALHL 284

QY 605 RKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDSDEGDDYCDGDEDEDLX 664

Db 285 RRYNDALLIHTVRAVDALAAALQDFYHREHVTKTQIL-----C----- 322

QY 665 KPLKLDETDRFLMTLFFENNKMRLKRLAENPEYENKLTKLRTIMEQYTRTEESARGIIF 724

Db 323 -----AERLLALFDDRKNELAHATHGP-ENPKLEMLEKILQRFSS-SSNSPRGIIF 373

QY 725 TKTRQSAYALSQWITENEKFAEVGVKHAHLIGAGHSSEFKPMQTONQKEVISKFTGKIN 784

Db 374 TRTRQSAHSLLLWLQOQGLQTVDIRAQLLIGAGNSSQSTHMTQRDQQEVQIKFQDGTIN 433

QY 785 LLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGVIEHE 844

Db 434 LLVATSAEEGLDIPHCNVVVRYGLLTNEISMVQARGRAWADQSVYAFVATEGSRRELKRE 493

QY 845 TVNDFREKMMYKAHCVQNMKPEEYAHKILELQMQSIMKMKTKRNIKHYKKNPSLIT 904

Db 494 LINEALETLMQAVAAVQKMDQAEYQAKIRDLQQAALTKRAAQAAQRENQRQOFPVEHVQ 553

QY 905 FLCKNCVSLACSGEDIHVIEKMHVNMTPEFKELY-IVRENKALQKXCADYQINGEIIICK 963

Db 554 LLCINCMVAVGHGSDLRKVEGTHHVNPNFNSYVNSRDPVINKVKDKWPGGVISCR 613

QY 964 -CGQAWGTMMVHKGLDLPCLKIRNFVVFVFNKSTKKQYKKWVLPITFPNLDYSECC 1019

Db 614 NCGEVWGLQMIYKSVKLPVLKVR--MLLETQGRQIAKWSRVFSPVDFDFLQHC 668

RESULT 13

Q9D2Z5 PRELIMINARY; PRT; 244 AA.

AC Q9D2Z5

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Adult male cecum cDNA, RIKEN full-length enriched library,

DE clone:913009C22 product:similar to melanoma differentiation

DE associated protein-5 (Fragment).

GN MDA5 OR 913009C22RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheraia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI TaxID=10090;

RN [1] -

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cecum;

RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,

RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,

RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,

RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,

RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,

RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,

RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,

RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,

RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,

RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,

RA Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cecum;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cecum;

RX MEDLINE=21085660; PubMed=11217851;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cecum;

RX MEDLINE=99279253; PubMed=10349636;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44(1999).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cecum;

RX MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cecum;

RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nishi K., Kitsunai T., Tashiro H., Itoh M.,

RA Konno H., Akiyama J., Nakamura S., Hazama M., Nishine T., Harada A.,

RA Sumi N., Ishii Y., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Yamamoto R., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

RA Fujiwaki S., Inoue K., Ozawa K., Tanaka T., Matsuura S., Kawai J.,

RA Yoneda Y., Ishikawa T., Inoue Y., Kira A., Hayashizaki Y.;

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format

RT sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).

DR EMBL; AK018602; BAB31303.2; --.

DR MGD; MGI:1918836; Mda5.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001650; Helicase_C.

DR Pfam; PF00271; helicase_C; 1.

FT NON TER 1

SQ SEQUENCE 244 AA; 28154 MW; 4D742F304832948C CRC64;

Query Match 21.1%; Score 1122; DB 11; Length 244;

Best Local Similarity 88.1%; Pred. No. 4.2e-55;

Matches 215; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 782 KINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGVI 841

Db 1 EINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVTSSGSGVT 60

QY 842 EHETVNDFREKMMYKAHCVQNMKPEEYAHKILELQMQSIMKMKTKRNIKHYKKNPS 901

Db 61 BREIVNDFREKMMYKAINRVQNMKPEEYAHKILELQVQSILEKKMKVERSIKQYNDNPS 120

QY 902 LITFLCKNCVSLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKXCADYQINGEII 961

Db 121 LITLLCKNCVSLVCSGENIHVIEKMHVNMTPEFKELYIVRENKALQKFCADYQTNGEII 180

QY 962 CKCGQAWGTMMVHKGLDLPCLKIRNFVVFVFNKSTKKQYKKWVLPITFPNLDYSECCLF 1021

Db 181 CKCGQAWGTMMVHKGLDLPCLKIRNFVVFVFNKSNPKKQYKKWVLPITFPDLDYSEYCLY 240

QY 1022 SDED 1025

Db 241 SDED 244

RESULT 14

O95786

ID O95786 PRELIMINARY; PRT; 925 AA.

Db 61 LLLSTLEKGVWHLGWTRFVEALRRGTGSPLAARYMNPBLTDLPSPSFENAHDEYLOLLNL 120
QY 121 LQPTLVDKLLVRDVLDKCMEELLTIEDRNRIAAAENNGNESGVRELLKRIVQKENWFSA 180
Db 121 LQPTLVDKLLVRDVLDKCMEELLTIEDRNRIAAAENNGNESGVRELLKRIVQKENWFSA 180
QY 181 FLNVLRTGNNELVOELTGSDCSESNAELENLSQVD 216
Db 181 FLNVLRTGNNELVOELTGSDCSESNAELENLSQVD 216

Search completed: June 2, 2004, 19:11:47
Job time : 53 secs